

Copyright (C) 1993 - 2002 Compugen Ltd.
GenCore version 5.1.3

CC This invention describes a novel use of bromelain protease (I) for
 CC inhibiting blood coagulation. (I) stimulate formation of plasmin, but
 CC inhibit both formation of fibrin and adhesion of thromocytes to
 CC endothelial cells. (I) are about 10 times more active than bromelain
 CC crude powder, i.e. comparable with streptokinase. AAYE1581 and AAY51582
 CC represent bromelain proteases isolated from pineapple (Ananas comosus)
 XX which are used in the method of the invention.

SQ Sequence 20 AA;

Query Match 100 %; Score 108; DB 21; Length 20;
 Best Local Similarity 100 %; Pred. No. 1.1e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VPOSIDWRDYGAVTSVKNQN 20
 Db 1 VPOSIDWRDYGAVTSVKNQN 20

RESULT 2

AAR47550

ID AAR47550 standard; protein; 213 AA.

XX AC AAR47550;

XX DT 18-JUL-1994 (first entry)

XX DE Stem bromelain protease.

XX KW Proteolytic enzyme; cysteine protease; diarrhoea; fluid secretion;

XX KW gastro-intestinal tract; GI tract; absorption.

XX OS Ananus comosus.

XX PN WO9400147-A.

XX PD 06-JAN-1994.

XX PF 30-JUN-1993; 93WO-GB01374.

XX PR 30-JUN-1992; 92GB-0013362.

XX PR 20-APR-1993; 93GB-0008164.

XX PR 25-JUN-1993; 93GB-0013189.

(CORT-) CORTECS LTD.

PS Disclosure; Page 11-13; 70pp; English.

XX PI Mynott TL;

XX DR WPI; 1994-025889/03.

XX PT Enzyme, esp. bromelain, contig. compsn. for control or diarrhoea -

PT esp. when caused by heat-labile toxin-producing E. coli, also

PT inhibits fluid secretion in the gut and promotes absorption

XX PS Disclosure; Page 11-13; 70pp; English.

XX SQ Sequence 213 AA;

XX DR WPI; 1994-025889/03.

XX PT Enzyme, esp. bromelain, contig. compsn. for control or diarrhoea -

PT esp. when caused by heat-labile toxin-producing E. coli, also

PT inhibits the secretory effect of the toxins and prevents attachment

CC of bacteria to the intestinal lining. It also inhibits secretion

CC caused by secretagogues, e.g. PGE₂, theophylline, etc. and has no

CC adverse effect on nutrient influx.

XX SQ Sequence 213 AA;

XX DR WPI; 1994-025889/03.

XX PT Enzyme, esp. bromelain, contig. compsn. for control or diarrhoea -

PT esp. when caused by heat-labile toxin-producing E. coli, also

PT inhibits fluid secretion in the gut and promotes absorption

XX PS Disclosure; Page 11-13; 70pp; English.

XX SQ Sequence 213 AA;

XX DR WPI; 1994-025889/03.

XX PT Enzyme, esp. bromelain, contig. compsn. for control or diarrhoea -

PT esp. when caused by heat-labile toxin-producing E. coli, also

PT inhibits fluid secretion in the gut and promotes absorption

XX PS Disclosure; Page 11-13; 70pp; English.

XX SQ Sequence 213 AA;

XX DR WPI; 1994-025889/03.

XX PT Enzyme, esp. bromelain, contig. compsn. for control or diarrhoea -

PT esp. when caused by heat-labile toxin-producing E. coli, also

PT inhibits fluid secretion in the gut and promotes absorption

Db 2 VPQSIDWRDYGAVTSVKNQN 21

RESULT 3

AAR68974

ID AAR68974 standard; protein; 213 AA.

XX AC AAR68974;

XX DT 09-AUG-1995 (first entry)

XX DE Pineapple stem bromelain protease.

XX KW Bromelain; Protease; proteolytic enzyme; cholera; diphtheria; cancer;

XX KW inflammation.

XX OS Ananas comosus.

XX PN WO9500163-A.

XX PD 05-JAN-1995.

XX PF 24-JUN-1994; 94WO-GB01368.

XX PR 25-JUN-1993; 93GB-0013188.

XX PA (CORT-) CORTECS LTD.

XX PI Mynott TL;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease - to treat conditions

PT dependent on cyclic nucleotide pathways e.g. cholera, diphtheria

PT or cancer

XX PS Claim 1; Page 35; 45pp; English.

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease - to treat conditions

PT dependent on cyclic nucleotide pathways e.g. cholera, diphtheria

PT or cancer

XX PS Claim 1; Page 35; 45pp; English.

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

Db 2 VPQSIDWRDYGAVTSVKNQN 21

RESULT 3

AAR68974

ID AAR68974 standard; protein; 213 AA.

XX AC AAR68974;

XX DT 09-AUG-1995 (first entry)

XX DE Pineapple stem bromelain protease.

XX KW Bromelain; Protease; proteolytic enzyme; cholera; diphtheria; cancer;

XX KW inflammation.

XX OS Ananas comosus.

XX PN WO9500163-A.

XX PD 05-JAN-1995.

XX PF 24-JUN-1994; 94WO-GB01368.

XX PR 25-JUN-1993; 93GB-0013188.

XX PA (CORT-) CORTECS LTD.

XX PI Mynott TL;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease - to treat conditions

PT dependent on cyclic nucleotide pathways e.g. cholera, diphtheria

PT or cancer

XX PS Claim 1; Page 35; 45pp; English.

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

XX SQ Sequence 213 AA;

XX DR WPI; 1995-051751/07.

XX PT Use of purified stem bromelain protease is used in the treatment of

CC prophylaxis of conditions which are dependent on cyclic nucleotide

CC and other intracellular and/or extracellular pathways. The diseases

CC include cancer, inflammation, atherosclerosis and bacterial

CC infections such as cholera, diphtheria and pertussis.

OS	Ananas comosus.
XX	WO20014253-A1.
PN	
XX	
PD	16-MAR-2000.
XX	
PP	24-AUG-1999; 99WO-GB02784.
XX	
PR	02-SEP-1998; 98GB-0019138.
XX	
PR	(CORT-) CORTECS UK LTD.
XX	
PA	(CORT-) CORTECS UK LTD.
XX	
PT	Myrott TL, Crossett B;
XX	
WPI	2000-25698/22.
XX	
DR	N-PSPB; AA293389.
XX	
PT	stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a pineapple stem -
PT	
PS	Claim 4; Page 58-59; 72pp; English.
XX	
CC	The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2.
CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours;
CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours;
CC	non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery), medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.
CC	
SQ	Sequence 311 AA;
Query Match	91.7%; Score 99; DB 21; Length 311;
Best Local Similarity	90.0%; Pred. No. 1e-08; Length 311;
Matches	18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 VPOSIDWDRDGAVTSTVNQN 20
Db	96 VPOSIDWDRDGAVNEVKNQN 115
RESULT 5	
AAV83107	
ID	AY83107 standard; Protein; 324 AA.
AC	AY83107;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	D385848 proteinase from pineapple stem.
XX	
DE	Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunosuppressant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; antimicrobial.
XX	
KW	Ananas comosus.
XX	
OS	WO20014253-A1.
XX	
PD	16-MAR-2000.
XX	
PP	24-AUG-1999; 99WO-GB02784.
XX	
PR	02-SEP-1998; 98GB-0019138.
XX	
PA	(CORT-) CORTECS UK LTD.
XX	
PT	Myrott TL, Crossett B;
XX	
DR	WPI; 2000-25698/22.
XX	
DR	N-PSPB; AA293391.
XX	
PT	stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a pineapple stem -
PT	
PS	Claim 4; Page 64-66; 72pp; English.
XX	
CC	The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2.
CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours;
CC	non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery), medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.
CC	
SQ	Sequence 324 AA;
Query Match	91.7%; Score 99; DB 21; Length 324;
Best Local Similarity	90.0%; Pred. No. 1e-08; Length 324;
Matches	18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 VPOSIDWDRDGAVTSTVNQN 20
Db	96 VPOSIDWDRDGAVNEVKNQN 115
RESULT 6	
AAV83104	
ID	AY83104.standard; Protein; 351 AA.
XX	
AC	AY83104;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	674 proteinase from pineapple stem.
XX	
DE	Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunosuppressant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; antimicrobial.
XX	
KW	Ananas comosus.
XX	
OS	WO20014253-A1.
XX	
PD	16-MAR-2000.
XX	

PF	24-AUG-1999;	99WO-GB02784.	XX	XX	(CORT-) CORTECS UK LTD.
PR	02-SEP-1998;	98GB-0019138.	XX	XX	Engwerda C, Mynott TL, Peek K;
PA	(CORT-) CORTECS UK LTD.		XX	XX	WPI: 1998-481194/41.
PA			XX	DR	New isolated component of bromelain - used for treating e.g. autoimmune diseases, transplant rejection, allergic reactions, toxic shock, apoptosis, parasitic or pathogen infections or cancer
PI			XX	PT	Disclosure; Page 24; 56pp; English.
PT			XX	PT	The patent discloses a fraction of bromelain, designated CCS, which
PT			XX	PT	(i) contains proteins having molecular weights of 15.07 kD, 25.85 kD and 27.45 kD (as determined by SDS-PAGE); (ii) has isoelectric points of 10.4 and 10.45, and (iii) is obtainable by: (a) dissolving bromelain in acetate buffer at pH 5.0; (b) separating the components by fast flow high performance chromatography on S-Sepharose (RTM), eluting with a linear gradient of 0 to 0.8M NaCl in acetate buffer, over 200 mL; (c) collecting the fraction corresponding to the final double peak of the column; and (d) isolating the protein from the fraction collected in (c). The CCS fraction can be used: (a) for modulating intracellular signalling pathways which control cell growth and proliferation; (b) for inhibiting the production of growth factors and cytokines by cells; (c) for reducing or preventing the activation of mitogen-activated protein (MAP) kinase pathway; (d) for reducing or preventing the activation of T cells; (e) as an immunosuppressive agent; (f) for blocking the production of growth factors and other cytokines; or (g) for the treatment or prevention of autoimmune diseases, graft or transplant rejection, allergic reactions, toxic shock, apoptosis, parasite or pathogen infections, or cancer. The CCS fraction is postulated to consist of the cysteine proteases ananain and comosain, together with various other components. The present sequence represents the N-terminal of comosain.
PS			XX	PS	Sequence 20 AA:
PS			XX	Query Match 89.8%; Score 97; DB 19; Length 20;	
PS	Claim 4; Page 55-57; 72pp; English.		XX	Best Local Similarity 94.7%; Pred No. 8.8e-10; Gaps 0; Gaps 0;	
CC	The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CX2.		XX	Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours;		XX	Qy 1 VPOSIDWRDYGAVTSVKNQ 19	
CC	non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery), medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.		Db 1 VPOSIDWRDYGAVTSVKNQ 19		
CC			XX	RESULT 8	
CC	Sequence 351 AA;		XX	AY83098	
CC	Query Match 91.7%; Score 99; DB 21; Length 351;		XX	ID AY83098 standard; Protein; 186 AA.	
CC	Best Local Similarity 90.0%; Pred. No. 1.1e-08; Matches 0; Mismatches 2; Indels 0; Gaps 0;		XX	AC AY83098;	
DE	Comosain N-terminal sequence.		XX	AC XX	
KW	Comosain; bromelain; CCS; autoimmune; immunosuppressant; graft rejection; antiallergic; antiapoptosis; toxic shock; antiparasitic; anticancer; antitumour.		XX	DT 04-JUL-2000 (first entry)	
KW	Ananas comosus.		XX	DB 569 proteinase from pineapple stem.	
OS	WO838291-A1.		XX	KW Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour;	
PN			XX	KW immunosimulant; immunodeficiency virus; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; anti-microbial.	
PD	03-SEP-1998.		XX	KW Ananas comosus.	
PD	25-FEB-1998;	98WO-GB00590.	XX	OS XX	
PR	25-MAR-1997;	97GB-0006119.	XX	PN WO200014253-A1.	
PR	25-FEB-1997;	97GB-0003827.	XX	PD 16-MAR-2000.	
PR	25-FEB-1997;	97GB-0003850.	XX	PP 24-AUG-1999;	
PR	28-FEB-1997;	97GB-0004252.	XX	PR 02-SEP-1998; 99WO-GB02784.	

PA (CORT-) CORTECS UK LTD.
 XX
 PI Myrott TL, Crossett B;
 XX DR WPI; 2000-256998/22.
 XX N-PSDB; AA293382.

PT Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a PT pineapple stem -
 XX
 PS Claim 4; Page 44-45; 72pp; English.

XX The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2. The proteinases can be used for treating or preventing cancer, e.g. solid tumours; non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant in the preparation of an immunosimilant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.

XX Sequence 186 AA;

Query	Match	Score	DB	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy	2 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0
Db	1 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0

XX

XX Sequence 186 AA;

Query	Match	Score	DB	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy	2 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0
Db	1 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0

XX

RESULT 9

AAV83101 AAY83101 standard; Protein; 186 AA.

XX AAV83101;

AC AAV83101;

XX DT 04-JUL-2000 (first entry)

DE 573 proteinase from pineapple stem.

XX Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunosimilant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; antimicrobial.

OS Ananas comosus.

XX

FR Key Location/Qualifiers

FT Misc-difference 20 /note= "Unidentified amino acid"

XX WO20014253-A1.

XX 16-MAR-2000.

PD 16-MAR-2000.

XX PR 24-AUG-1999; 99WO-GB02784.

XX 02-SEP-1998; 98GB-0019138.

XX (CORT-) CORTECS UK LTD.

XX

XX PA (CORT-) CORTECS UK LTD.
 XX
 PI Myrott TL, Crossett B;
 XX DR WPI; 2000-256998/22.
 XX N-PSDB; AA293385.

PT Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a PT pineapple stem -
 XX
 PS Claim 4; Page 50-51; 72pp; English.

XX The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2. The proteinases can be used for treating or preventing cancer, e.g. solid tumours; non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunosimilant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.

XX Sequence 186 AA;

Query	Match	Score	DB	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy	2 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0
Db	1 POSIDWRDYGAVNEVKNON	95	21	186	89.5%			0	0	0

XX

RESULT 10

AAV83099 AAV83099 standard; Protein; 187 AA.

XX AAV83099;

AC AAV83099;

XX DT 04-JUL-2000 (first entry)

DE 571 proteinase from pineapple stem.

XX Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunosimilant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; antimicrobial.

OS Ananas comosus.

XX

FR PN WO20014253-A1.

XX 16-MAR-2000.

XX PR 24-AUG-1999; 99WO-GB02784.

XX 02-SEP-1998; 98GB-0019138.

XX (CORT-) CORTECS UK LTD.

XX

PI	Myntott TL, Crossett B;	XX	Myntott TL, Crossett B;
XX		PI	
DR	WPI; 2000-256998/22.	XX	
DR	N-PSDB; AA293383.	DR	WPI; 2000-256998/22.
XX		N-PSDB; AA293384.	
PT	Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production by nucleic acid sequences isolated from a pineapple stem -	XX	Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a pineapple stem -
PS	Claim 4; Page 46-47; 72pp; English.	XX	Claim 4; Page 48-49; 72pp; English.
XX		XX	
CC	The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2.	CC	The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CCX2.
CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours; non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.	CC	The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours; non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.
SQ	Sequence 187 AA;	SQ	Sequence 188 AA;
Query	Match 88.0%; Score 95; DB 21; Length 187;	Query	Match 88.0%; Score 95; DB 21; Length 188;
Best Local Similarity 89.5%; Pred. No. 2.7e-08;	Best Local Similarity 89.5%; Pred. No. 2.7e-08;	Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	2 POSIDWRDYGAVTSVKNQN 20	Qy	2 PQSIDWRDYGAVTSVKNQN 20
Db	1 POSIDWRDYGAVNEVKNQN 19	Db	1 PQSIDWRDYGAVNEVKNQN 19
RESULT 1.1		RESULT 1.2	
ID	AAY83100 standard; Protein: 190 AA.	ID	AAY83106 standard; Protein: 190 AA.
XX		XX	
AC	AAY83100;	AC	AAY83106;
XX		XX	
DT	04-JUL-2000 (first entry)	DT	04-JUL-2000 (first entry)
XX		XX	
DE	572 proteinase from pineapple stem.	DE	D14059 proteinase from pineapple stem.
XX		XX	
KW	Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunostimulant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; generic abnormality; diabetes; nitric oxide; NO; antimicrobial.	KW	Proteinase; pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunostimulant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; generic abnormality; diabetes; nitric oxide; NO; antimicrobial.
XX		XX	
OS	Ananas comosus.	OS	Ananas comosus.
XX		XX	
FH	Location/Qualifiers	FN	WO200014253-A1.
FT	Misc-difference 72	FT	02-SEP-1998;
FT	/note= "Unidentified amino acid"	FT	(CORT-) CORTECS UK LTD.
PN	W0200014253-A1.	PN	16-MAR-2000.
XX		XX	XX
PD	16-MAR-2000.	PD	16-MAR-2000.
XX		XX	
PF	24-AUG-1999;	PF	24-AUG-1999;
XX	99WO-GB02784.	XX	99WO-GB02784.
PF	24-AUG-1999;	PR	02-SEP-1998;
XX	98GB-0019138.	XX	98GB-0019138.
PR	02-SEP-1998;	PA	(CORT-) CORTECS UK LTD.
XX		XX	
PA	(CORT-) CORTECS UK LTD.	PI	Myntott TL, Crossett B;
XX		XX	

DR WPI; 2000-256998/22.

DR N-PSDB; AAZ33390.

XX Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a pineapple stem -

XX Claim 4; Page 61; 72pp; English.

CC The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CxK2. The proteinases can be used in human or veterinary medicine and can be used for treating or preventing cancer, e.g. solid tumours; non-solid tumours and virus induced tumours. The proteinases can also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with drugs such as steroids, cyclosporin and cyclophosphamide), protein loss (e.g. in diarrhea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinase can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.

XX Sequence 190 AA;

Query Match Score 88.0%; Score 95; DB 21; Length 190; Best Local Similarity 89.5%; Pred. No. 2.8e-08; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy ||||||| POSIDWRDYGAVTSVKNQ 20
Db 1 POSIDWRDYGAVNEVKNQ 19

SQ

RESULT 13
AAW0801 ID AAW0801 standard; peptide; 20 AA.

AC XX

AC AAW0801;

XX DT 24-NOV-1998 (first entry)

XX DE Ananain N-terminal sequence.

XX KW Ananain; bromelain; CGS; autoimmune; immunosuppressant;

XX KW graft rejection; antileptic; antiapoptosis; toxic shock;

XX KW anti-parasitic; antipathogenic; anticancer; antitumour.

XX OS Ananas comosus.

XX PN WO9338291-A1.

XX PD 03-SEP-1998.

XX PF 25-FEB-1998; 98WO-GB00590.

XX PR 25-MAR-1997; 97GB-0006119.

XX PR 25-FEB-1997; 97GB-0003827.

XX PR 25-FEB-1997; 97GB-0003850.

XX PR 28-FEB-1997; 97GB-0004252.

PA (CORT-) CORTECS UK LTD.
XX PI Engwerda, C., Mynott, T.L., Peek, K;
XX WPI; 1998-481194/41.

PT New isolated component of bromelain - used for treating e.g.

PT autoimmune diseases, transplant rejection, allergic reactions, toxic shock, apoptosis, parasite or pathogen infections or cancer

XX Disclosure; Page 24; 56pp; English.

CC The patent discloses a fraction of bromelain, designated CCS, which (i) contains proteins having molecular weights of 15.07 kD, 25.85 kD and 27.45 kD (as determined by SDS-PAGE); (ii) has isoelectric points of 10.4 and 10.45, and (iii) is obtainable by: (a) dissolving bromelain in acetate buffer at pH 5.0; (b) separating the components by fast flow high performance chromatography on S-Sepharose (RTM), eluting with a linear gradient of 0 to 0.8M NaCl in acetate buffer over 300 ml; (c) collecting the fraction corresponding to the final double peak off the column; and (d) isolating the protein from the fraction collected in (c). The CCS fraction can be used: (a) for modulating intracellular signalling pathways which control cell growth and proliferation; (b) for inhibiting the production of growth factors and cytokines by cells; (c) for reducing or preventing the activation of mitogen activated protein (MAP) kinase pathway; (d) for reducing or preventing the activation of T cells; (e) as an immunosuppressive agent; (f) for blocking the production of growth factors and other cytokines; or (g) for the treatment or prevention of autoimmune diseases, graft or transplant rejection, allergic reactions, toxic shock, apoptosis, parasite or pathogen infections, or cancer. The CCS fraction is postulated to consist of the cysteine proteases ananain and comosain, together with various other components. The present sequence represents the N-terminal of ananain.

XX Sequence 20 AA;

Query Match Score 86.1%; Score 93; DB 19; Length 20; Best Local Similarity 94.7%; Pred. No. 4.4e-09; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDYGAVTSVKNQ 19
Db 1 VPOSIDWRDGSAGVTSVKNQ 19

SQ

RESULT 14
AYY51582 ID AYY51582 standard; peptide; 20 AA.

AC XX

AC AYY51582;

XX DT 18-MAY-2000 (first entry)

XX DE Pineapple bromelain protease peptide fragment from fraction 5.

XX KW Bromelain; pineapple; protease; blood coagulation; inhibitor; plasmin.

XX OS Ananas comosus.

XX PN WO200033729-A1.

XX PD 27-JAN-2000.

XX PF 15-JUL-1998; 98WO-EP04406.

XX PR 15-JUL-1998; 98WO-EP04406.

XX PA (URSA-) URSAPHARM ARZNEIMITTEL GMBH.

XX PI Maurer, R., Eckert, K., Grabowska, E., Eschmann, K;

XX DR WPI; 2000-182338/16.

XX PT Composition for inhibiting blood coagulation more efficiently than crude bromelain powder comprises bromelain proteases -

XX PS Claim 8; Page 16; 24pp; German.

XX CC This invention describes a novel use of bromelain proteases (I) for inhibiting blood coagulation. (I) stimulate formation of plasmin, but

CC inhibit both formation of fibrin and adhesion of thrombocytes to CC endothelial cells. (I) are about 10 times more active than bromelain CC crude powder, i.e. comparable with streptokinase. AAY51581 and AAY51582 CC represent bromelain proteases isolated from Pineapple (Ananas comosus) which are used in the method of the invention.

XX Sequence 20 AA;

Query Match 86.1%; Score 93; DB 21; Length 21;

Best Local Similarity 94.7%; Pred. No. 4.4e-09; Mismatches 0;

Matches 18; Conservative 0; Indels 0; Gaps 0;

QY 1 VPOSIDWRDYGAVTSVKNQ 19

Db 1 VPOSIDWRDGSAGAVTSVKNQ 19

Search completed: November 21, 2002, 16:08:39

Job time : 32 secs

RESULT 15
ID AAY84395 standard; peptide; 21 AA.
XX
AC AAY84395;
XX
DT 12-JUL-2000 (first entry)
XX
DE N-terminal peptide from a 26 kDa protein of bromelain.
XX
KW Bromelain; CCS/P2; ERK-2 phosphorylation; interleukin-2;
KW CD4+ T cell proliferation; mitogen activated protein kinase; MAP kinase;
KW growth factor; cytokine; cancer; parasite; pathogen infection;
KW immunosuppressant; autoimmune disease; graft rejection;
KW transplant rejection; allergic reaction; toxic shock; apoptosis.
XX
OS Ananas comosus.
XX
PH Key
PT Misc-difference 9
FT /note= "not specified"
FT 10
FT Misc-difference 10
FT /note= "not specified"
XX
PN WO200014254-A2.
XX
PD 16-MAR-2000.
XX
PF 24-AUG-1999; 93WO-GB02732.
XX
PR 02-SEP-1998; 98GB-0019137.
XX
PA (CORT-) CORTecs UK LTD.
XX
PI Mynt TL, Peek K, Tyremann DR;
XX
DR 2000-256999/22.
XX

Bromelain fractions useful in the prevention and treatment of cancer, parasites and pathogen infections and autoimmune diseases comprise proteins with biological activity -
XX
PS Claim 5; Page 52; 69pp; English.
XX
CC The present sequence represents a N-terminal sequence from a 24 kDa protein (as measured by SDS-PAGE). The protein is derived from a fraction of bromelain, designated CCS/P2, comprising proteins of molecular weights 23.9 kDa and 26.6 kDa when measured by SDS-PAGE or 23.459 kDa and 23.6576 kDa when measured by mass spectrometry. The fraction blocks ERK-2 (serine/threonine kinase) phosphorylation and MAP kinase cascade. They also block interleukin-2 (IL-2) production and CD4+ T cell proliferation. The bromelain fractions and proteins are used in medicine as they inhibit the mitogen activated protein (MAP) kinase pathway, activate T cells and produce growth factors and cytokines. They can be used for treating cancer, parasites or pathogen infections. They can be used as an immunosuppressant for treating or preventing autoimmune diseases, graft

CC or transplant rejection by a host, allergic reactions, toxic shock or CC apoptosis. The fractions and proteins may also be used to prepare agents CC for the treatment of the above diseases and conditions.
XX
SQ Sequence 21 AA;

Query Match 86.1%; Score 93; DB 21; Length 21;

Best Local Similarity 90.0%; Pred. No. 4.e-09; Mismatches 0;

Matches 18; Conservative 0; Indels 0; Gaps 0;

QY 1 VPOSIDWRDYGAVTSVKNQ 20

Db 1 VPOSIDWRXGAVTSVKNQ 20

Search completed: November 21, 2002, 16:08:39

Job time : 32 secs

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 15:59:31 ; Search time 31 Seconds
 (without alignments)
 85.968 Million cell updates/sec

Searched: US-09-674-738-2
 Perfect score: 105
 Sequence: 1 VPQSIDWDRGAVTIVSRKNG 20

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:
 Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
 Listing first 45 summaries

Database : A_GeneSeq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1980 DAT: *
 2: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1981 DAT: *
 3: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1982 DAT: *
 4: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1983 DAT: *
 5: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1985 DAT: *
 6: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1986 DAT: *
 7: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1987 DAT: *
 8: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1988 DAT: *
 9: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1989 DAT: *
 10: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1990 DAT: *
 11: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1991 DAT: *
 12: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1992 DAT: *
 13: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1993 DAT: *
 14: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1994 DAT: *
 15: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1995 DAT: *
 16: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1996 DAT: *
 17: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1997 DAT: *
 18: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1998 DAT: *
 19: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA1999 DAT: *
 20: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA2000 DAT: *
 21: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA2001 DAT: *
 22: /SIDS2/gcgdata/geneseq/geneseqp-embj/AA2002 DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	105	100.0	20 19	AAWB0801
2	105	100.0	20 21	AYV1582
3	105	100.0	28 21	AYV4389
4	105	100.0	35 21	AYV84391
5	100	95.2	21 21	AYV84394
6	94	89.5	20 19	AAWB0802
7	93	88.6	20 21	AYV51581
8	93	88.6	213 15	ARK7550
9	93	88.6	213 16	ARK8974
10	88	83.8	21 21	AYV84395

11 85 81.0 21 21 AAY84396
 12 84 80.0 311 21 AAY83105
 13 84 80.0 324 21 AAY83107
 14 84 80.0 351 21 AAY83104
 15 83 79.0 377 20 AAW89356
 16 82 78.1 127 22 AAB65793
 17 82 78.1 148 22 AAB65784
 18 82 78.1 278 22 AAB65782
 19 81 77.1 258 21 AAG44392
 20 81 77.1 277 21 AAG30215
 21 81 77.1 282 21 AAG30214
 22 81 77.1 294 21 AAG21091
 23 81 77.1 297 21 AAG44291
 24 81 77.1 300 21 AAG44290
 25 81 77.1 312 21 ARG0213
 26 81 77.1 361 21 AAG50423
 27 81 77.1 362 21 AAG50422
 28 81 77.1 368 21 AAG25448
 29 81 77.1 369 22 AAB65770
 30 81 77.1 376 22 AAB65770
 31 80 76.2 129 21 AAY3098
 32 80 76.2 166 21 AAY3098
 33 80 76.2 186 21 AAY83101
 34 80 76.2 187 21 AAY83099
 35 80 76.2 188 21 AAY83100
 36 80 76.2 190 21 AAY83106
 37 80 76.2 442 18 AAW27412
 38 80 76.2 443 18 AAW27441
 39 79 75.2 15 21 AAY84397
 40 79 75.2 18 21 AAW84985
 41 79 75.2 20 21 AAY84393
 42 79 75.2 21 21 AAY84387
 43 79 75.2 21 21 AYV84387
 44 79 75.2 231 19 AAW72388
 45 79 75.2 345 22 AAB85128

ALIGNMENTS

RESULT 1
 ID AAW80801
 ID AAW80801 standard; peptide; 20 AA.
 XX
 AC AAW80801;
 XX
 DT 24-NOV-1998 (first entry)
 XX
 DE Ananain N-terminal sequence.
 XX
 KW Ananain; bromelain; CCS; autoimmune; immunosuppressant; graft rejection; antiallergic; antiapoptosis; toxic shock; antiparasitic; antipathogenic; anticancer; antitumour.
 KW
 OS Ananas comosus.
 XX
 PN WO98388291-A1.
 XX
 PR 03-SEP-1998.
 XX
 PF 25-FEB-1998; 98WO-GB00590.
 XX
 PR 25-MAR-1997; 97GB-0006119.
 PR 25-FEB-1997; 97GB-0003827.
 PR 25-FEB-1997; 97GB-0003850.
 PR 28-FEB-1997; 97GB-0004252.
 XX
 (CORT-) CORTECS UK LTD.
 XX
 PI Engwerda C., Myntott TL, Peek K;
 XX
 DR WPI; 1998-481194/41.



PT New isolated component of bromelain - used for treating e.g.
 PT autoimmune diseases, transplant rejection, allergic reactions,
 PT toxic shock, apoptosis, parasite or pathogen infections or cancer
 Disclosure; Page 24: 56pp; English.

The patent discloses a fraction of bromelain, designated CCS, which
 CC (i) contains proteins having molecular weights of 15.07 kD, 25.85 kD
 CC and 27.45 kD (as determined by SDS-PAGE); (ii) has isoelectric points
 CC of 10.4 and 10.45, and (iii) is obtainable by: (a) dissolving bromelain
 CC in acetate buffer at pH 5.0; (b) separating the components by fast flow
 CC high performance chromatography on S-Sepharose (RTM), eluting with a
 CC linear gradient of 0 to 0.8M NaCl in acetate buffer over 300 ml; (c)
 CC collecting the fraction corresponding to the final double peak off the
 CC column; and (d) isolating the protein from the fraction collected in (c).
 CC The CCS fraction can be used: (a) for modulating intracellular signalling
 CC pathways which control cell growth and proliferation; (b) for inhibiting
 CC the production of growth factors and cytokines by cells; (c) for reducing
 CC or preventing the activation of mitogen activated protein (MAP) kinase
 CC (e) for reducing or preventing the activation of T cells;
 CC (f) as an immunosuppressive agent; (g) for blocking the production of
 CC growth factors and other cytokines; or (g) for the treatment or
 prevention of autoimmune diseases, graft or transplant rejection,
 CC allergic reactions, toxic shock, apoptosis, parasite or pathogen
 CC infections, or cancer. The CCS fraction is postulated to consist of other
 CC the cysteine proteases ananain and comosain, together with various other
 CC components. The present sequence represents the N-terminal of ananain.
 XX

Sequence 20 AA;
 Query Match 100.0%; Score 105; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.6e-10;
 Matches 20; Conservative 0; Misnatches 0; Indels 0; Gaps 0;

QY 1 VPOSIDWRSGAVTSVNGQ 20
 Db 1 VPQSIDWRSGAVTSVNGQ 20

RESULT 2
 AAY51582 standard; peptide; 20 AA.

XX
 DE Pineapple bromelain protease peptide fragment from fraction 5.
 XX
 KW Bromelain; pineapple; protease; blood coagulation; inhibitor; plasmin.
 XX
 OS Ananas comosus.
 XX
 PN WO200003729-A1.
 XX
 PD 27-JAN-2000.
 XX
 PP 15-JUL-1998; 98WO-EP044406.
 XX
 PR 15-JUL-1998; 98WO-EP044406.
 XX
 PA (URSA-) URSAPHARM ARZNEIMITTEL GMBH.
 XX
 PI Maurer R, Eckert K, Grabowska E, Eschmann K;
 XX
 WPI; 2000-182338/16.

XX
 PT Composition for inhibiting blood coagulation more efficiently than
 PT crude bromelain powder comprises bromelain proteases -
 XX
 PS Claim 8; Page 16; 24pp; German.
 XX
 CC This invention describes a novel use of bromelain proteases (I) for
 CC

PT stimulating formation of fibrin and adhesion of thromocytes to
 PT endothelial cells. (I) are about 10 times more active than bromelain
 PT crude power, i.e. comparable with streptokinase. AAY51581 and AAY51582
 PS represent bromelain proteases isolated from pineapple (Ananas comosus)
 XX
 CC which are used in the method of the invention.

Sequence 20 AA;
 Query Match 100.0%; Score 105; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.6e-10;
 Matches 20; Conservative 0; Misnatches 0; Indels 0; Gaps 0;
 QY 1 VPOSIDWRSGAVTSVNGQ 20
 Db 1 VPQSIDWRSGAVTSVNGQ 20

RESULT 3
 AAY84389 standard; peptide; 28 AA.

ID AAY84389
 XX
 AC AAY84389;
 XX
 DT 12-JUL-2000 (first entry)

XX
 DE N-terminal peptide from a fraction of bromelain designated CCS/P2.
 XX
 Bromelain; CCS/P2; ERK-2 phosphorylation; interleukin-2;
 KW CD4+ T cell proliferation; mitogen activated protein kinase; MAP kinase;
 KW growth factor; cytokine; cancer; parasite; pathogen infection;
 KW immunosuppressant; autoimmune disease; graft rejection;
 KW transplant rejection; allergic reaction; toxic shock; apoptosis.

OS Ananas comosus
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 22 /note= "not specified"
 FT Misc-difference 25 /note= "not specified"
 FT Misc-difference 26 /note= "not specified"
 FT Misc-difference 27 /note= "not specified"
 PN WO200014254-A2.
 XX
 PA (CORTES) CORTES UK LTD.
 XX
 PD 16-MAR-2000.
 XX
 PR 99WO-GB02792.
 XX
 PR 02-SEP-1998; 98GB-0019137.
 XX
 DR WPI; 2000-256999/22.

XX
 PT Bromelain fractions useful in the prevention and treatment of cancer,
 PT parasites and pathogen infections and autoimmune diseases comprise
 PT proteins with biological activity -
 XX
 PS Claim 1; Page 51; 69pp; English.

XX
 CC The present sequence represents a N-terminal sequence from a fraction
 CC of bromelain, designated CCS/P2, comprising proteins of molecular weights
 CC 23.9 kDa and 26.6 kDa when measured by SDS-PAGE or 23.45 kDa and 22.6576
 CC kDa when measured by mass spectrometry. The peptides block ERK-2
 CC (serine/threonine kinase) phosphorylation and MAP kinase cascade. They
 CC also block interleukin-2 (IL-2) production and CD4+ T cell proliferation.
 CC The bromelain fractions and proteins are used in medicine as they inhibit
 CC the mitogen activated protein (MAP) kinase pathway, activate T cells and
 CC produce growth factors and cytokines. They can be used for treating

immunosuppressant for treating or preventing autoimmune diseases, graft or transplant rejection by a host, allergic reactions, toxic shock or apoptosis. The fractions and proteins may also be used to prepare agents for the treatment of the above diseases and conditions.

CC infections, or cancer. The CCS fraction is postulated to consist of other CC the cysteine proteases ananain and comosain, together with various other CC components. The present sequence represents the N-terminal of comosain.
XX

Sequence 21 AA:

Ananas comosus	WO9838291-Al.	03-SEP-1998.	25-FEB-1998;	98WO-GB00590.
			25-MAR-1997;	97GB-0006119.
			25-FEB-1997;	97GB-0003827.
			97GB-0003850.	
			28-FEB-1997;	97GB-0004252.

(CORT-) CORTECS UK LTD.
Engwerda C, Mynott TL,
Peek K,
WDT 1000 123104/1

WP1 / 1994-481194-41.
New isolated component of bromelain - used for treating e.g., autoimmune diseases, transplant rejection, allergic reactions, toxic shock, apoptosis, parasite or pathogen infections or cancer
Disclosure: Page 24; 56pp; English.

The patent discloses a fraction of bromelain, designated CCS, which (ii) contains proteins having molecular weights of 15.07 kD, 25.85 and 27.45 kD (as determined by SDS-PAGE), (iii) has isoelectric points of 10.4 and 10.45, and (iv) is obtainable by: (a) dissolving bromelain in acetate buffer at pH 5.0; (b) separating the components by high performance chromatography on S-Sepharose (RTM), eluting with linear gradient of 0 to 0.8M NaCl in acetate buffer over 300 ml; collecting the fraction corresponding to the final double peak of the column; and (d) isolating the protein from the fraction collected. The CCS fraction can be used: (a) for modulating intracellular signal pathways which control cell growth and proliferation; (b) for inhibiting the production of growth factors and cytokines by cells; (c) for activating T cells or preventing the activation of mitogen-activated protein (MAP) kinase pathway; (d) for reducing or preventing the activation of T cells (e) as an immunosuppressive agent; (f) for blocking the production of growth factors and other cytokines; or (g) for the treatment or prevention of autoimmune diseases, graft or transplant rejection, allergic reactions, toxic shock, anaphylaxis or parasitic or pathogen infections.

CC infections, or cancer. The CCS fraction is postulated to consist of other CC the cysteine proteases ananain and comosain, together with various other CC components. The present sequence represents the N-terminal of comosain.
XX

```

Sequence 21 AA:
Query Match 95.2%; Score 100; DB 21;
Best Local Similarity 95.0%; Pred. No. 1.1e-09;
Matches 19; Conservative 0; Mismatches 1;
Indels 0; Gaps 0

1 VPOSIDWDRSGAVTSKNOG 20
1 VPOSIDWDRSGAVTSKNOG 20
1 VPOSIDWDRSGAVTSKNOG 20

```

SULT 6
W80802 AAW80802 standard, peptide; 20 AA.

AC PAYS1581;
 XX DT 18-MAY-2000 (First entry)
 XX DE Pineapple bromelain protease peptide fragment from fraction 4.
 XX KW Bromelain; pineapple; protease; blood coagulation; inhibitor; plasmin.
 XX OS Ananas comosus.
 OS PN WO200003739-A1.

<i>Ananas comosus</i> .	
WO9818291-A1.	
03-SEP-1998.	
25-FEB-1998:	
25-MAR-1997;	97GB-0006119-
25-FEB-1997;	97GB-0003827-
25-FEB-1997;	97GB-0003850-
28-FEB-1997;	97GB-0004252-

(CORT-) CORTECS UK L
Engwerda C, Mynott
SPT 12000 121124/11

CC infections, or cancer. The CCS fraction is postulated to consist of other CC the cysteine proteases ananain and comosain, together with various other CC components. The present sequence represents the N-terminal of comosain.
XX

RESULT 7
AAV51581 standard; peptide; 20 AA.
ID AAV51581
XX

AC PAYS1581;
 XX DT 18-MAY-2000 (First entry)
 XX DE Pineapple bromelain protease peptide fragment from fraction 4.
 XX KW Bromelain; pineapple; protease; blood coagulation; inhibitor; plasmin.
 XX OS Ananas comosus.
 OS PN WO200003739-A1.

XX Composition for inhibiting blood coagulation more efficiently than
PT crude bromelain powder comprises bromelain proteases -
XX
PS Claim 6; Page 15; 24pp; German.
xx

This invention describes a novel use of bromelain proteases (I) for inhibiting blood coagulation. (I) stimulate formation of plasmin, but inhibit both formation of fibrin and adhesion of thrombocytes to endothelial cells. (I) are about 10 times more active than bromelain crude powder, i.e. comparable with streptokinase. AY5151 and AY51582 represent bromelain proteases isolated from pineapple (*Ananas comosus*)

CC which are used in the method of the invention.
XX SQ Sequence 20 AA;
SQ Sequence March 20, 1986, Serial No. 032, 000, Inventor, 20.

```

QY      1 VPQSDDWRSGAVTSVRNQ 19
      Best Local Similarity 99.7%; score 32;  db 2.; length 20;
      Query Match 99.7%; score 32;  db 2.; length 20;
      Best Local Matches 18;  Conservative 0;  Mismatches -1;  Indels 0;  Gaps

```

RESULT	8
Db	1 VPQSIDWRDYGAIVTSVKNQ 19
AAR47550	
ID	AAR47550 standard; protein; 213 AA.
XX	
AC	AAR47550;
XX	

DT 18-JUL-1994 (First entry)
 XX
 DE Stem bromelain protease.
 XX
 KW Proteolytic enzyme; cysteine protease; diarrhoea; fluid secretion;
 KW gastro-intestinal tract; GI tract; absorption.
 OS Ananas comosus.
 XX
 PN WO9400147-A.
 PD 06-JAN-1994.
 XX
 PP 30-JUN-1993; 93WO-GB01374.
 XX
 PR 30-JUN-1992; 92GB-001382.
 PR 20-APR-1993; 93GB-0008164.
 PR 25-JUN-1993; 93GB-0013189.
 XX
 (CORT-) CORTECS LTD.
 PA
 XX
 PI Myhott TL;
 XX
 DR WPI; 1994-025889/03.
 XX
 PT Enzyme, esp. bromelain, contg. compn. for control or diarrhoea -
 PT esp. when caused by heat-labile toxin-producing E. coli, also
 PT inhibits fluid secretion in the gut and promotes absorption
 XX
 PS Disclosure; Page 11-13; 70pp; English.

XX
 CC The sequence is that of stem bromelain protease, a mixture of
 CC various moieties derived from the stem of the pineapple. The
 protein is a cysteine protease used to treat infectious diarrhoea
 CC caused by heat labile or heat stable toxins, enterotoxigenic E. coli
 CC and Vibrio cholerae. The enzyme reduces toxin binding activity,
 CC inhibits the secretory effect of the toxins and prevents attachment
 CC of bacteria to the intestinal lining. It also inhibits secretion
 CC caused by secretagogues, e.g. PGE2, theophylline, etc. and has no
 CC adverse effect on nutrient influx.

XX
 SQ Sequence 213 AA;

Query Match 88.6%; Score 93; DB 15; Length 213;
 Best Local Similarity 94.7%; Pred. No. 2.4e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 VPOSIDWRSQAVTSVKHQ 19
 Db 2 VPOSIDWRSQAVTSVKHQ 20

RESULT 10
 AAY84395
 ID AAY84395 standard; peptide; 21 AA.
 XX
 AC AAY84395;
 XX
 DT 12-JUL-2000 (first entry)
 XX
 DE N-terminal peptide from a 26 kDa protein of bromelain.
 XX
 KW Bromelain; ccsP2; ERK-2 phosphorylation; interleukin-2;
 KW growth factor; cytokine; cancer; parasite; pathogen infection;
 KW immunosuppressant; autoimmune disease; graft rejection;
 KW transplant rejection; allergic reaction; toxic shock; apoptosis.
 XX
 OS Ananas comosus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 9 /note= "not specified"
 FT Misc-difference 10 /note= "not specified"
 FT
 XX
 PN WO200014254-A2.
 XX
 PD 16-MAR-2000.
 XX
 PR 24-AUG-1999; 99WO-GB02792.
 XX
 PR 02-SEP-1998; 98GB-0019137.
 XX
 PA (CORT-) CORTECS UK LTD.
 XX
 PI Myhott TL, Peek K, Tyreman DR;
 XX
 DR WPI; 2000-256999-22.
 XX
 PT Bromelain fractions useful in the prevention and treatment of cancer,
 PT parasites and pathogen infections and autoimmune diseases comprise
 XX

RESULT 9
 AAR68974
 ID AAR68974 standard; protein; 213 AA.
 XX
 AC AAR68974;
 XX
 DT 09-AUG-1995 (first entry)
 XX
 DE Pineapple stem bromelain protease.
 XX
 KW Bromelain; protease; proteolytic enzyme; cholera; diphtheria; cancer;
 KW inflammation.
 OS Ananas comosus.
 XX
 PN WO9500165-A.
 XX
 PD 05-JAN-1995.
 XX
 PR 24-JUN-1994; 94WO-GB01368.
 XX

PS Claim 5; Page 52; 69pp; English.
 XX The present sequence represents a N-terminal sequence from a 24 kDa protein (as measured by SDS-PAGE). The protein is derived from a fraction of bromelain, designated CCS/P2, comprising proteins of molecular weights 23.9 kDa and 26.6 kDa when measured by SDS-PAGE or 23.459 kDa and 23.6576 kDa when measured by mass spectrometry. The fraction blocks ERK-2 (serine/threonine kinase phosphorylation and MAP kinase cascade. They also block interleukin-2 (IL-2) production and CD4+ T cell proliferation, (serine/threonine kinase cascade. They can be used as an immunosuppressant for treating or preventing autoimmune diseases, graft or transplant rejection by a host, allergic reactions, toxic shock or apoptosis. The fractions and proteins may also be used to prepare agents for the treatment of the above diseases and conditions.

XX Sequence 21 AA;

Query Match Score 88; DB 21; Length 21;
 Best Local Similarity 89.5%; Pred. No. 1.le-07; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPQSIDWRDGAIVSVKQ 19

Db 1 VPQSIDWRDGAIVSVKQ 19

RESULT 11

ID AAY84396 standard; peptide; 21 AA.

XX AAY84396;

XX DT 12-JUL-2000 (first entry)

DE N-terminal peptide from a 24 kDa protein of bromelain.

XX Bromelain; CCS/P2; ERK-2 phosphorylation, interleukin-2; KW growth factor; cytokine; cancer; parasite; pathogen infection; KW immunosuppressant; autoimmune disease; graft rejection; KW transplant rejection; allergic reaction; toxic shock; apoptosis.

XX Ananas comosus

XX Location/Qualifiers

FH Key Misc-difference 17 /note= "not specified"

FT Misc-difference 19 /note= "not specified"

FT WO2000014254-A2.

XX 16-MAR-2000.

XX 24-AUG-1999; 99WO-GB02784.

XX 02-SEP-1998; 98GB-0019138.

XX (CORT-) CORTECS UK LTD.

XX Mynot TL, Peak K, Tyreman DR;

XX WPI; 2000-256999/22.

XX 02-SEP-1998; 98GB-0019137.

XX Bromelain fractions useful in the prevention and treatment of cancer, parasites and pathogen infections and autoimmune diseases comprise proteins with biological activity -

PS Claim 6; Page 53; 69pp; English.

CC The present sequence represents a N-terminal sequence from a 24 kDa protein (as measured by SDS-PAGE). The protein is derived from a fraction of bromelain, designated CCS/P2, comprising proteins of molecular weights 23.9 kDa and 26.6 kDa when measured by SDS-PAGE or 23.459 kDa and 23.6576 kDa when measured by mass spectrometry. The fraction blocks ERK-2 (serine/threonine kinase phosphorylation and MAP kinase cascade. They also block interleukin-2 (IL-2) production and CD4+ T cell proliferation, (serine/threonine kinase cascade. They can be used as an immunosuppressant for treating or preventing autoimmune diseases, graft or transplant rejection by a host, allergic reactions, toxic shock or apoptosis. The fractions and proteins may also be used to prepare agents for the treatment of the above diseases and conditions.

XX Sequence 21 AA;

Query Match Score 85; DB 21; Length 21;
 Best Local Similarity 95.0%; Pred. No. 3.5e-07; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGAIVSVKQ 20

Db 1 VPOSIDWRDGAIVSVKQ 20

RESULT 12

ID AAY83105 standard; Protein; 311 AA.

XX AAY83105;

XX AC AAY83105;

XX DT 04-JUL-2000 (first entry)

XX DE 710 proteinase from pineapple stem.

XX Proteinase; Pineapple; treatment; prevention; prophylaxis; tumour; solid tumour; non-solid tumour; virus induced tumour; immunostimulant; immunodeficiency; HIV; human immunodeficiency virus; malaria; malnutrition; trauma; burn; surgery; genetic abnormality; diabetes; nitric oxide; NO; antimicrobial.

XX Ananas comosus

XX OS WO2000014253-A1.

XX PN WO2000014253-A1.

XX PD 16-MAR-2000.

XX PP 24-AUG-1999; 99WO-GB02784.

XX PR 02-SEP-1998; 98GB-0019138.

XX (CORT-) CORTECS UK LTD.

XX PI Mynot TL, Crossett B;

XX DR WPI; 2000-256998/22.

XX DR N-PSDB; A4/293389.

XX PT Stem bromelain-like proteinases useful for the prevention and treatment of cancer and diseases or conditions which respond to increased nitric oxide production are encoded by nucleic acid sequences isolated from a pineapple stem -

XX PS Claim 4; Page 58-59; 72pp; English.

XX PT The sequences encoding proteinases from pineapple stem have the same activity as a stem bromelain-like proteinase, CXc2. CC The proteinases can be used in human or veterinary medicine and CC non-solid tumours and virus induced tumours. The proteinases can

also be used as an immunostimulant or in the preparation of an immunostimulant to treat immunodeficiencies arising from malnutrition, infection (e.g. HIV and malaria), tumours (e.g. lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical treatment (e.g. with cyclophosphamide), protein loss (e.g. in diarrhoea and burns), genetic abnormalities (e.g. in combined immunodeficiency patients lacking B or T cells), diabetes and old age. The proteinate can also be used in the treatment or prevention of diseases or conditions which respond to increased nitric oxide (NO) production, or as an antimicrobial agent or the component of an antimicrobial agent.

CC
CC immunostimulant to treat immunodeficiencies arising from
CC malnutrition, infection (e.g. HIV and malaria), tumours (e.g.
CC lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical
CC treatment (e.g. with drugs such as steroids, cyclosporin and
CC cyclophosphamide), protein loss (e.g. in diarrhoea and burns),
CC genetic abnormalities (e.g. in combined immunodeficiency patients
CC lacking B or T cells), diabetes and old age. The proteinate can
CC also be used in the treatment or prevention of diseases or
CC conditions which respond to increased nitric oxide (NO) production,
CC or as an antimicrobial agent or the component of an antimicrobial
CC agent.

CC
XX Sequence 311 AA;

SQ

Query Match 80.0%; Score 84; DB 21; Length 311;
Best Local Similarity 84.2%; Pred. No. 1.2e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRISGAVNSVVKQ 19
Db 96 VPOSIDWRISGAVNSVVKQ 114

RESULT 13
AY83107
ID AY83107 standard; Protein: 324 AA.

XX AC AY83107;

XX DT 04-JUL-2000 (first entry)

DE D185848 proteinase from pineapple stem.

XX KW proteinase; pineapple; treatment; prevention; prophylaxis; tumour;
KW solid tumour; non-solid tumour; virus induced tumour;
KW immunostimulant; immunodeficiency; HIV;
KW human immunodeficiency virus; malaria; malnutrition; trauma; burn;
KW surgery; genetic abnormality; diabetes; nitric oxide; NO;
KW antimicrobial.
OS Ananas comosus.

XX OS Ananas comosus.

PN WO200014253-A1.

XX XX

PD 16-MAR-2000.

XX PP 24-AUG-1999; 99WO-GB02784.

XX PR 02-SEP-1998; 98GB-0019138.

(CORT-) CORTECS UK LTD.

XX PI Myrott TL, Crossett B;

XX DR WPI; 2000-25698/22.

XX DR N-PsDB; AA293391.

PT Stem bromelain-like proteinases useful for the prevention and treatment
PT of cancer and diseases or conditions which respond to increased nitric
PT oxide production are encoded by nucleic acid sequences isolated from a
PT pineapple stem.

PT Claim 4; Page 64-66; 72pp; English.

CC The sequences encoding proteinases from pineapple stem have
CC the same activity as a stem bromelain-like proteinase, CCX2.
CC The proteinases can be used in human or veterinary medicine and
CC can be used for treating or preventing cancer, e.g. solid tumours;
CC non-solid tumours and virus induced tumours. The proteinases can
CC also be used as an immunostimulant or in the preparation of an
CC immunostimulant to treat immunodeficiencies arising from
CC malnutrition, infection (e.g. HIV and malaria), tumours (e.g.
CC lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical
CC treatment (e.g. with drugs such as steroids, Cyclosporin and
CC cyclophosphamide), protein loss (e.g. in diarrhea and burns),

CC lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical
CC treatment (e.g. with drugs such as steroids, cyclosporin and
CC cyclophosphamide), protein loss (e.g. in diarrhoea and burns),
CC genetic abnormalities (e.g. in combined immunodeficiency patients
CC lacking B or T cells), diabetes and old age. The proteinate can
CC also be used in the treatment or prevention of diseases or
CC conditions which respond to increased nitric oxide (NO) production,
CC or as an antimicrobial agent or the component of an antimicrobial
CC agent.

CC
XX Sequence 324 AA;

SQ

Query Match 80.0%; Score 84; DB 21; Length 324;
Best Local Similarity 84.2%; Pred. No. 1.2e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRISGAVNSVVKQ 19
Db 96 VPOSIDWRISGAVNSVVKQ 114

RESULT 14
AY83104
ID AY83104 standard; Protein: 351 AA.

XX AC AY83104;

XX DT 04-JUL-2000 (first entry)

DE 674 proteinase from pineapple stem.

XX KW proteinase; pineapple; treatment; prevention; prophylaxis; tumour;

KW solid tumour; non-solid tumour; virus induced tumour;

KW immunostimulant; immunodeficiency; HIV;

KW human immunodeficiency virus; malaria; malnutrition; trauma; burn;

KW surgery; genetic abnormality; diabetes; nitric oxide; NO;

KW antimicrobial.

XX OS Ananas comosus.

XX PN WO200014253-A1.

XX DD 16-MAR-2000.

XX PF 24-AUG-1999; 99WO-GB02784.

XX PR 02-SEP-1998; 98GB-0019138.

XX PA (CORT-) CORTECS UK LTD.

XX PI Myrott TL, Crossett B;

XX DR WPI; 2000-25698/22.

XX DR N-PsDB; AA293398.

XX PT Stem bromelain-like proteinases useful for the prevention and treatment
PT of cancer and diseases or conditions which respond to increased nitric
PT oxide production are encoded by nucleic acid sequences isolated from a
PT pineapple stem.

XX PS Claim 4; Page 55-57; 72pp; English.

CC The sequences encoding proteinases from pineapple stem have
CC the same activity as a stem bromelain-like proteinase, CCX2.
CC The proteinases can be used in human or veterinary medicine and
CC can be used for treating or preventing cancer, e.g. solid tumours;
CC non-solid tumours and virus induced tumours. The proteinases can
CC also be used as an immunostimulant or in the preparation of an
CC immunostimulant to treat immunodeficiencies arising from
CC malnutrition, infection (e.g. HIV and malaria), tumours (e.g.
CC lymphoid, myeloma), trauma (e.g. burns, wounds, surgery) medical
CC treatment (e.g. with drugs such as steroids, Cyclosporin and
CC cyclophosphamide), protein loss (e.g. in diarrhea and burns),

CC genetic abnormalities (e.g. in combined immunodeficiency patients
 CC lacking B or T cells), diabetes and old age. The proteinase can
 CC also be used in the treatment or prevention of diseases or
 CC conditions which respond to increased nitric oxide (NO) production,
 CC or as an antimicrobial agent or the component of an antimicrobial
 CC agent.

XX SQ Sequence 351 AA;

Query Match 80.0%; Score 84; DB 21; Length 351;
 Best Local Similarity 84.2%; Pred. No. 1.3e-05;Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 ID AAW9556 standard; Protein; 377 AA.

AC AAW9556;

XX AC

XX DT

16-MAR-1999

(First entry)

XX DE

Triticum sp. cysteine proteinase #1.

XX KW

Triticum; wheat seed; cysteine proteinase; gluten; baking,

XX OS

Triticum sp.

XX PN

JP10327686-A.

XX PD

15-DEC-1998.

XX PP

27-MAR-1998;

98JP-0098140.

XX PR

31-MAR-1997;

97JP-0114946.

XX PA

(SHOS) SHOWA SANGYO CO.

XX DR

WPI; 1999-109255/10.

N-PSDB; AAV82456.

XX PT

New DNA coding cysteine proteinase originating from wheat seed -

PT useful for improving gluten for use in bakery process

XX PS

Claim 1; Page 15-16; 29pp; Japanese.

XX

The present sequence represents a cysteine proteinase isolated from
 CC wheat seed (Triticum sp.). The cysteine proteinase is useful for
 CC improving gluten for use in the bakery process.

XX SQ Sequence 377 AA;

Query Match 79.0%; Score 83; DB 20; Length 377;
 Best Local Similarity 70.0%; Pred. No. 2.1e-05;Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 ID VPOSIDWRDSSGAVTSVKNQ 20

Qy 1

Db 138 LPRSVDMRKQGAVTGVKNQ 157

Copyright (c) 1993 - 2002 Compugen Ltd.	Gencore version 5.1.3			
Run on:	November 21, 2002, 15:59:48 ; Search time 7.5 Seconds			
Title:	US-09-674-738-2			
Perfect score:	105			
Sequence:	1 VFQSIDWRDGSAGTYSVKNQG 20			
Scoring table:	BLOSUM62			
Searched:	Gapop 10.0 , Gapext 0.5			
Total number of hits satisfying chosen parameters:	112892			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
Post-processing:	Minimum Match 0% Listing first 45 summaries			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
Database :	SwissProt_40;*			
Query	%			
Result No.	Score	Match Length	DB ID	Description
1	105	100.0	345	1 ANAN_ANACO
2	93	88.6	212	1 BROM_ANACO
3	80	80.0	352	1 CATT_CARRA
4	83	79.0	334	1 CATT_PIG
5	82	78.1	43	1 CC3_GARCN
6	82	78.1	43	1 CCA_GARCN
7	82	78.1	334	1 CATT_MOUSE
8	81	77.1	319	1 CATT_SCHNA
9	81	77.1	361	1 CATT_ARATH
10	81	77.1	368	1 RD19_ARATH
11	80	76.2	221	1 GPI_ZINOF
12	80	76.2	471	1 ORYB_ORYEA
13	79	75.5	221	1 CATT_RAT
14	79	75.2	334	1 CATT_RAT
15	79	75.2	345	1 PAPA_CARRA
16	79	75.2	354	1 CYS1_LEITL
17	79	75.2	354	1 LCPA_LEITL
18	78	74.3	341	1 CATT_DROME
19	78	74.3	362	1 C1SP_PHATO
20	78	74.3	362	1 CYSP_VIGMO
21	78	74.3	443	1 LCPB_LEITL
22	78	74.3	443	1 CYS2_LEITL
23	77	73.3	333	1 CATT_HUMAN
24	76	72.4	348	1 PAPA_CARRA
25	76	72.4	371	1 CYS1_HORTU
26	76	72.4	373	1 CYS2_HORTU
27	74	70.5	343	1 CYS1_DICCI
28	74	70.5	462	1 RD21_ARATH
29	73	69.5	330	1 CATS_RAT
30	73	69.5	363	1 CYSP_PEA
31	73	69.5	395	1 CATT_BRUCA
32	72	68.6	218	1 CATT_SHEEP
33	72	68.6	1 CATT_CHICK	
				ALIGNMENTS
				RESULTS
				ANAN_ANACO
				SEQUENCE OF 123-338.
				TISSUE=Stem; STANDARD;
				PRT; 345 AA.
				ID : ANAN_ANACO
				AC : P80884; 02293;
				DT : 01-NOV-1997 (Rel. 35, Created)
				DT : 16-OCT-2001 (Rel. 40, Last sequence update)
				DT : 15-JUN-2002 (Rel. 41, Last annotation update)
				DB : Ananain precursor (EC 3.4.22.31).
				GN : ANI.
				RC : Robertson, C. E., Goodnough, P. W.;
				RA : Biocat, J. 327:1199-202(1997).
				RT : "Cloning and expression of ananain gene from pineapple.";
				RL : Submitted (Nov-1997) to the EMBL/GenBank/DDBJ databases.
				RN : [2]
				RN : NCBI_TAXID=6115;
				RN : [1]
				RN : SEQUENCE FROM N.A.
				RR : TISSUE=Stem; TISSUE=Stem;
				RC : STRAIN=cv. Smooth Cayenne;
				RC : Robertson, C. E., Goodnough, P. W.;
				RC : Bromelaceae; Ananas.
				OC : OX : 34395; arabidopsis
				OC : P43916; arabidopsis
				OC : P82474; zingiber of
				OC : P25777; oryza sativ
				OC : P82473; zingiber of
				OC : P07154; rattus norv
				OC : P07849; carica papa
				CC : P35591; leishmania
				CC : P25759; leishmania
				CC : Q95029; drosophila
				CC : P25803; phaeocetes v
				CC : P12412; vigna mungo
				CC : P36000; leishmania
				CC : Q05094; leishmania
				CC : P07711; homo sapien
				CC : P05994; carica papa
				CC : P25549; hordeum vul
				CC : P2520; hordeum vul
				CC : P04988; dictyosteli
				CC : P43207; arabidopsis
				CC : Q02165; rattus norv
				CC : P25804; pisum sativ
				CC : Q17413; brugia paha
				CC : Q10991; ovis aries
				CC : P09648; gallus gall
				P25774; homo sapien
				P15242; rattus norv
				P25975; bos taurus
				Q10716; zea mays (m
				P25776; oryza sativa
				P32954; carica cana
				P10056; carica papa
				P25326; bos taurus
				P25251; brassica napus
				P20721; lycopersico
				P25779; mus musculus
				P55097; mus musculus

RESULT 6
 CC4_CARCN STANDARD; PRT; 43 AA.
 ID CC4_CARCN
 AC P3257; (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteine proteinase IV (EC 3.4.22.-) (CC-IV) (Fragment).
 OS Carica candaarcensis.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 Eurosids II; Brassicales; Caricaceae; Carica.
 NCBI_TaxID=29731; [1]
 RN SEQUENCE.
 RP TISSUE=LateX;
 RC MEDLINE=4030669; PubMed=8216902;
 RA Waireavans V., Jaziri M., van Beeumen J., Schnek A.G.,
 RA Kleinschmidt T., Loosz Y.;
 RA "Isolation and preliminary characterization of the cysteine-
 proteinases from the latex of Carica candaarcensis Hook.";
 RT Proteinases from the latex of Carica candaarcensis Hook.;
 RL Biol. Chem. Hoppe-Seyler 374:501-506 (1993).
 CC ~ SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 DR HSSP; P14080; YAL.
 DR MEROPS; C01_0PA; --.
 DR InterPro; IPR000168; Peptidase C1.
 DR InterPro; IPR000169; SHprc_ac5ite.
 DR Pfam; PF00112; Peptidase_C1_1.
 DR ProDom; PD000158; Peptidase C1_1.
 DR PROSITE; PS00639; THIOL_Protease_HIS; PARTIAL.
 DR PROSITE; PS00640; THIOL_Protease ASN; PARTIAL.
 DR PROSITE; PS00139; THIOL_Protease_CYS; 1.
 KW Hydrolase; Thiol protease.
 FT ACT SITE 25 25 BY SIMILARITY.
 FT NON_TER 43 43
 SQ SEQUENCE 43 AA; F4D5945A9386E291 CRC64;
 Query Match 78.1%; Score 82; DB 1; Length 43;
 Best Local Similarity 78.9%; Pred. No. 3.2e-07;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 PQSIDWRSQAVTSVKNGQ 20
 Db 2 PESTIDWRSQAVTPVRNGQ 20

RESULT 7
 CATL_MOUSE STANDARD; PRT; 334 AA.
 ID CATL_MOUSE
 AC P06797; (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP).
 OS Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090; [1]
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88016849; PubMed=2689328;

RA Troen B.R., Gal S., Gotteman M.M.;
 RT "Sequence and expression of the CDNA for MEP (major excreted
 protein), transformation-regulated secreted cathepsin.";
 RL Biochem. J. 246:731-735 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88213715; PubMed=2835398;
 RA Joseph L.J., Chang L.C., Stamenkovich D., Sukhatme V.P.;
 "Complete nucleotide and deduced amino acid sequences of human and

RT murine preprocathepsin L. An abundant transcript induced by
 transformation of fibroblasts.";
 RL J. Clin. Invest. 81:1621-1629 (1988).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87033683; PubMed=3533924;
 RA Portnoy D.A., Erickson A.H., Kochan J.J., Ravetch J.V., Unkeless J.C.;
 RT "Cloning and characterization of a mouse cysteine proteinase.";
 RL J. Biol. Chem. 261:14697-14703 (1986).
 RN [4]
 RN SEQUENCE FROM N.A., AND CARBOHYDRATE-LINKAGE SITE.
 RX TISSUE=Liver;
 RC MEDLINE=91112761; PubMed=2275556;
 RA Stearns N.A., Dong J., Pan J.X., Brenner D.A., Sahagian G.G.;
 RT "Comparison of cathepsin L synthesised by normal and transformed
 cells at the gene, message, protein, and oligosaccharide levels.";
 RL Arch. Biochem. Biophys. 283:447-457 (1990).
 RN [5]
 RN SEQUENCE OF B9-300 FROM N.A.
 RX STRAUN=BNL;
 RX MEDLINE=86271744; PubMed=3755373;
 RA Denhardt D.T., Hamilton R.T., Parfitt C.L.J., Edwards D.R.,
 RA Pierre R.S., Waterhouse P., Nilsson-Hamilton M.;
 RT "Close relationship of the major excreted protein of transformed
 murine fibroblasts to thiol-dependent cathepsins.";
 RL Cancer Res. 46:4530-4536 (1986).
 CC -!- FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN
 CC LYROSOMES
 CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain. As
 CC compared to cathepsin B, cathepsin L exhibits higher activity
 CC towards protein substrates, but has little activity on Z-Arg-Arg-
 CC NHMeC, and no peptidyl-dipeptidase activity
 CC -!- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE
 CC BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to licenses@isb-sib.ch).

CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to licenses@isb-sib.ch).

CC ---
 DR EMBL; X05086; CAA29470.1; --.
 DR EMBL; J05583; AAA7445.1; --.
 DR EMBL; M20455; AAA9984.1; --.
 DR EMBL; X04392; CAA27980.1; --.
 DR PIR; S01177; KHM5L.
 DR PIR; S13890; S13890.
 DR HSSP; P0771; ICJL.
 DR MBR05; C01.032; --.
 DR MGD; MGI:88564; Cts1.
 DR InterPro; IPR00068; Peptidase_C1.
 DR InterPro; IPR000169; SHprc_acsite.
 DR Pfam; PF00112; Peptidase_C1_1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; P000158; Peptidase_C1_1.
 DR PROSITE; PS00139; THIOL_Protease_CYS; 1.
 DR PROSITE; PS00639; THIOL_Protease_HIS; 1.
 DR PROSITE; PS00640; THIOL_Protease ASN; 1.
 KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 113
 FT CHAIN 114 288
 FT CHAIN 291 334
 FT ACT SITE 138 138
 FT ACT SITE 276 276
 FT ACT SITE 300 300
 FT DISUFD 135 178
 FT DISUFD 169 211
 FT DISUFD 269 322
 ACTIVATION PEPTIDE.
 CATHESIN L, HEAVY CHAIN.
 CATHESIN L, LIGHT CHAIN.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 INTERCHAIN (BY SIMILARITY).

FT	CARBOHYD	221	221	N-LINKED (GLCNAC, . . .).
FT	CONFLICT	58	58	M -> I (IN REF. 2); Best Local Similarity 65.0%; Pred. No. 4.26-06; Mismatches 4; Indels 0; Gaps 0;
FT	CONFLICT	177	177	G -> R (IN REF. 3); CRC64; Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
SQ	SEQUENCE	334 AA:	37547 MW: FEE747043307AD98	
QY	1 VPGSIDWRSQGAVTIVNQG 20			
Db	114 IPKPSVWDREKGCVTPVKNG 133			
RESULT 8				
CATL_SCHMA				
ID	CATL_SCHMA			
STANDARD:				
PRT:	319 AA.			
AC	026534;			
DT	15-JUN-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DR	Cathepsin L precursor (EC 3.4.22.15) (SMCL1).			
GN	CatL			
OS	Schistosoma mansoni (Blood fluke).			
OC	Bukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeida;			
OC	Schistosomatidae; Schistosomatidae; Schistosoma.			
RN	[1] NCBI_TaxID:6183;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Puerto Rican;			
RX	MEDLINE=9514043; Pubmed=7838171;			
RA	Smith A.M., Dalton J.P., Clough K.A., Kilbane C.L., Harrop S.A.,			
RA	Hole N., Brindley P.J.,			
RT	"Adult Schistosoma mansoni express cathepsin L proteinase activity.";			
RL	Mol. Biochem. Parasitol. 67:11-19(1994).			
CC	-!- FUNCTION: MAY BE CRUCIAL FOR METABOLISM OF HOST HEMOGLOBIN.			
CC	-!- CATALYTIC ACTIVITY: Specificity close to that of pepsin. As compared to cathepsin B, cathepsin L exhibits higher activity on Z-Arg-Arg-towards protein substrates, but has little activity on Z-Arg-Arg-NH ₂ , and no peptidyl-dipeptidase activity.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isp-sib.ch).			
CC				
CC	EMBL: U07345; AAC6485.1; -			
DR	HSSP: 06911; IFR0.			
DR	MEROPS: Col1.0.18; -			
DR	InterPro: IPR000668; Peptidase_C1.			
DR	InterPro: IPR000169; Sprotease_C1.			
DR	InterPro: IPR000669; Sprotease_C1.			
DR	pfam: PF0012; Peptidase_C1; 1.			
DR	PRINTS: PRO0058; Papain.			
DR	PRODOM: PD000158; Papain.			
DR	PROSITE: PS00139; THIOL_Protease_CYS; 1.			
DR	PROSITE: PS00639; THIOL_Protease_HIS; 1.			
DR	PROSITE: PS0064; THIOL_Protease ASN; 1.			
KW	Hydrolyase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.			
FT	Signal?			
FT	ACTIVATION PEPTIDE (POTENTIAL).			
FT	CHAIN 105 319 CATHESIN_L.			
FT	ACT_SITE 129 129 BY SIMILARITY.			
FT	ACT_SITE 265 286 BY SIMILARITY.			
FT	ACT_SITE 286 286 BY SIMILARITY.			
FT	DISULFID 126 167 BY SIMILARITY.			
FT	DISULFID 160 200 BY SIMILARITY.			
FT	DISULFID 258 307 BY SIMILARITY.			
SQ	SEQUENCE 319 AA: 36136 MW: 9CFB68A3FL193479 CRC64;			
FT	MECP2; Col1.0.22; -			
DR	InterPro: IPR000668; Peptidase_C1.			
DR	InterPro: IPR000169; Sprotease_C1.			
DR	pfam: PF0012; Peptidase_C1; 1.			
DR	PRINTS: PRO0058; Papain.			
DR	PRODOM: PD000158; Papain.			
DR	PROSITE: PS00139; THIOL_Protease_CYS; 1.			
DR	PROSITE: PS00639; THIOL_Protease_HIS; 1.			
DR	PROSITE: PS0064; THIOL_Protease ASN; 1.			
KW	Hydrolyse; Thiol protease; Zymogen; Glycoprotein; Signal.			
FT	SIGNAL?			

RESULT 10

ID RD19_ARATH	STANDARD;	PRT;	368 AA.
AC P43296;			
DT 01-NOV-1995 (Rel. 32, Created)			
DT 01-NOV-1995 (Rel. 32, Last sequence update)			
DT 15-JUN-2002 (Rel. 41, Last annotation update)			
DE Cysteine proteinase RD19 precursor (EC 3.4.22.-).			
GN RD19 OR At4g39090 OR F19A2.190.			
OS Arabidopsis thaliana (Mouse-ear cress).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyt; eudicots; Rosidae;			
OC eurosidae II; Brassicales; Brassicaceae; Arabidopsis.			
OX NCBI_TaxID=3702;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=cv. Columbia;			
RX MEDLINE=3314960; PubMed=8325504;			
RA Koiwumi M., Yamashuchi-Shinozaki K., Tsuji H., Shirozaki K.;			
RT "Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in Arabidopsis thaliana.";			
RL Gene 129:175-182(1993).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=cv. Columbia;			
RX MEDLINE=2008486; PubMed=10617198;			
RA Mayer K.F.X., Schueler C., Wambutt R., Murphy G., Volckaert G.,			
RA Pohl T., Duesterhoeft A., Strelkova W., Entian K.-D., Terryn N.,			
RA Harris B., Anzorge W., Brandt P., Grivell L.A., Rieger M.,			
RA Weichselgarther M., de Simone V., Obermaier B., Mache R., Mueller M.,			
RA Kreis M., Deiseny M., Puigdomenech P., Watson M., Schmidtkeini T.,			
RA Reichert B., Porretelle D., Perez-Alonso M., Bourry M., Bancroft I.,			
RA Vos P., Hoheisel J., Zimmerman W., Wedler H., Ridley P.,			
RA Langham S.-A., McCullagh B., Billingham L., Robben J.,			
RA Van der Schueren J., Gromponez B., Chuang Y.-J., Vandenburghe F.,			
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,			
RA Weissenegger T., Borle G., Ramsperger U., Hilbert H., Braun M.,			
RA Holzer E., Brandt A., Peeters S., van Staveren M., Dirksen W.,			
RA Klein-Lankhorst R., Rose M., Haut J., Koetter P.,			
RA Bernheiser S., Hempel S., Felipauch M., Lambeth S., Van den Daele H.,			
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,			
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-AlLEN S.,			
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McClay K., Mayes R.,			
RA Petett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,			
RA Borrova D., Bloecker H., Schaefer M., Grimm M., Lohner T.-H.,			
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Aier S.,			
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Hezzl A.,			
RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,			
RA Massenet O., Quigley F., Clabaud G., Muendlein R., Feilzer R.,			
RA Schrabl S., Hillier R., Schmidt W., Aubourg S., Lecharny A., Casacuberta E.,			
RA Cherdor F., Cooke R., Berger C., Montfort A., Torres J., Torroba J.,			
RA Gibbons T., Weber N., Vandebol M., Bargues M., Terol J., Johnson S., Tacon D., Jesse T.,			
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,			

ACTIVATION PEPTIDE (POTENTIAL)

PROPEP 24 131

PROTEINASE A494.

PROTEINASE PROBABLE CYSTEINE PROTEINASE

PROTEINASE BY SIMILARITY.

PROTEINASE ACT SITE 132 361

PROTEINASE ACT SITE 156 156

PROTEINASE ACT SITE 299 299

PROTEINASE ACT SITE 326 326

PROTEINASE DISULFID 153 203

PROTEINASE DISULFID 187 237

PROTEINASE DISULFID 293 347

PROTEINASE CONFLICT 49 49

PROTEINASE SEQUENCE 361 AA; 39819 MW; D092CEB5654642DD CRC64;

Query Match Best Local Similarity 77.0%; Score 81; Length 361; Pred. No. 4.8e-06; Mismatches 2; Gaps 0;

Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VPQSIDWRSIDSGAVTSYRKNQG 20

Dy 132 LPBEEFDWDRGAVTPYKRNQG 151

CC -!- INDUCTION: BY HIGH SALT CONDITIONS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR D13042; BAA02373_1; -

CC DR EMBL; AL035679; CAB3829_1; -

CC DR HSSP; P0711_1CUL.

CC DR MEROPS; C01_022; -

CC DR InterPro; IPR000169; Peptidase_C1.

CC DR PF00112; Peptidase_C1; 1.

CC DR PRINTS; PR00705; PAPAIN.

CC DR ProDom; PD000158; Peptidase_C1; 1.

CC DR PROSITE; PS00639; THIOL_Protease_CYS; 1.

CC DR PROSITE; PS00640; THIOL_Protease ASN; 1.

CC KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.

CC FT SIGNAL; 1 23 POTENTIAL.

CC FT PROPEP 24 134 ACTIVATION PEPTIDE (POTENTIAL).

CC FT CHAIN 135 368 CYSTEINE PROTEINASE RD19A.

CC FT ACT SITE 159 159 BY SIMILARITY.

CC FT ACT SITE 302 302 BY SIMILARITY.

CC FT ACT SITE 329 329 BY SIMILARITY.

CC FT DISUFID 156 206 BY SIMILARITY.

CC FT DISUFID 194 240 BY SIMILARITY.

CC FT CARBOHYD 296 350 BY SIMILARITY.

CC FT CARBOHYD 253 350 N-LINKED (GUCNAC). . . (POTENTIAL).

CC SQ SEQUENCE 368 AA; 40418 MW; 9d64CP35E07F519D CRC64;

Query Match Best Local Similarity 77.1%; Score 81; DB 1; Length 368; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRSIDSGAVTSYRKNQG 20

Db 135 LPEDFDWRDGHAVTPYKRNQG 154

RESULT 11

GPI1_ZINOF	STANDARD;	PRT;	221 AA.
ID GPI1_ZINOF			
AC P82474;			
DT 30-MAY-2000 (Rel. 39, Created)			
DT 30-MAY-2000 (Rel. 39, Last sequence update)			
DT 15-JUN-2002 (Rel. 41, Last annotation update)			
DT Cysteine Proteinase GP-11 (EC 3.4.22.-).			

Zingiber officinale (Ginger).
 OS Zingiber officinale
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Zingiberaceae;
 OC Zingiber
 OX NEBI_TaxID:94328;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Root;
 RX MEDLINE=20156257; PubMed=10691991;
 RA Choi K.H.; Laursen R.A.;
 RT "Amino-acid sequence and glycan structures of cysteine proteases with proline specificity from ginger rhizome Zingiber officinale.";
 RL Eur. J. Biochem. 267:1516-1526(2000).
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage with a proline residue at P2.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.
 HSSP: P00785; ZACT:
 Glycosidase, P22474; -.
 InterPro: IPR000668; Peptidase CL.
 InterPro: IPR000169; SHPr0t_acsite.
 Pfam: PF00112; Peptidase_CL_1.
 PRINTS: PRO00705; PAPAIN.
 Prodrom: PD000158; Peptidase CL_1.
 DR PROSTB: PS00139; THIOL_PROTEASE_CYS_1.
 DR PROS1; PS00639; THIOL_PROTEASE_HIS_1.
 DR PROS1; PS00640; THIOL_PROTEASE ASN; FALSE_NEG.
 KW Hydrolase; Thiol protease; Glycoprotein.
 FT ACT_SITE 27 27 BY SIMILARITY.
 FT ACT_SITE 161 161 BY SIMILARITY.
 FT DISULFID 24 65 BY SIMILARITY.
 FT DISULFID 58 98 BY SIMILARITY.
 FT DISULFID 155 206 BY SIMILARITY.
 FT CARBOHYD 99 99 /FTID=CAR_000190. . .).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .).
 FT /FTID=CAR_000200. . .).
 FT SEQUENCE 221 AA; 23922 MW; 909A312Bd8632D42 CRC64;
 Query Match 76.2%; Score 80; DB 1; Length 221;
 Best Local Similarity 70.0%; Pred. No. 4.2e-06;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 OX 1 VPGSIDWRDSGAVTSVKNGQ 20
 Db 3 LPNSIDWRENGAVVPVKNGQ 22

RESULT 12
 ORB_ORYSA STANDARD; PRT; 471 AA.
 AC P25777;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oryzain beta chain precursor (EC 3.4.22.-).
 Oryzain beta chain precursor (EC 3.4.22.-).
 Oryza sativa (Rice).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrholioideae; Oryzeae; Oryza.
 OX NEBI_TaxID=4530;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=CV_Nipponbare; TISSUE=Seed;
 RX MEDLINE=91358494; PubMed=1885617;
 RA Watanabe H.; Abe K.; Emori Y.; Hosoyama H.; Arai S.;
 RT Molecular cloning and gibberellin-induced expression of multiple cysteine proteases of rice seeds (oryzains).";
 RL J. Biol. Chem. 266:16897-16902(1991).
 CC -!- TISSUE_SPECIFICITY: EXPRESSED ONLY IN SEEDS.
 CC -!- INDUCTION: BY GIBBERELIC ACID (GAI).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
 CC ---
 DR InterPro; IPR00118; Granulin.
 DR InterPro; IPR000668; Peptidase CL.
 DR InterPro; IPR000169; SHPr0t_acsite.
 DR Pfam; PF00112; Peptidase_CL_1.
 DR Pfam; PF00396; Granulin; 1.
 DR PRINTS; PRO00705; PAPAIN.
 DR Prodrom; PD000188; Peptidase_CL_1.
 DR SMART; SM02277; GRAN; 1.
 DR PROS1; PS00640; THIOL_PROTEASE ASN; 1.
 KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 21 BY POTENTIAL.
 FT PROPEP 22 139 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 140 471 ORYZAIN BETA CHAIN.
 FT ACT_SITE 164 164 BY SIMILARITY.
 FT ACT_SITE 301 301 BY SIMILARITY.
 FT ACT_SITE 321 321 BY SIMILARITY.
 FT DISULFID 161 204 BY SIMILARITY.
 FT DISULFID 195 237 BY SIMILARITY.
 FT DISULFID 295 346 BY SIMILARITY.
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 471 AA; 50505 MW; 0BF39D3395C6B1D CRC64;

Query Match 76.2%; Score 80; DB 1; Length 471;
 Best Local Similarity 65.0%; Pred. No. 9.5e-06;
 Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OX 1 VPGSIDWRDSGAVTSVKNGQ 20
 Db 140 LPNSIDWRENGAVVPVKNGQ 159

RESULT 13
 GPI_ZINOF STANDARD; PRT; 221 AA.
 ID GPI_ZINOF
 AC P22473;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteine proteinase GP-I (EC 3.4.22.-).
 OS Zingiber officinale (Ginger).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Zingiberaceae;
 OX Zingiber
 OX NEBI_TaxID=94328;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Root;
 RX MEDLINE=20156257; PubMed=10691991;
 RA Choi K.H.; Laursen R.A.;
 RT "Amino-acid sequence and glycan structures of cysteine proteases with proline specificity from ginger rhizome Zingiber officinale.";
 RL Eur. J. Biochem. 267:1516-1526(2000).
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage with a proline residue at P2.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.
 CC -!- CAUTION: THE AUTHORS REGARD THE SEQUENCE AS TENTATIVE, AS THEY BELIEVE THAT IT MAY HAVE BEEN CONTAMINATED BY A HOMOLOGOUS PROTEIN.
 CC HSSP; P00785; ZACT.

DR	InterPro; IPR000668; Peptidase_C1.
DR	InterPro; IPR000169; SHprt_acsite.
DR	Pfam; PF00112; Peptidase_C1; 1.
PRINTS	PR00105; PAPAIN.
DR	ProDom; PD000158; Peptidase C1_1.
DR	PROSITE; PS00139; THIOL_PROTEASE_HIS; FALSE_NEG.
DR	PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
DR	PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
KW	Hydrolase; Thiol protease; Glycoprotein.
FT	ACT SITE 27 BY SIMILARITY.
FT	ACT SITE 161 BY SIMILARITY.
FT	DISULFIDE 24 BY SIMILARITY.
FT	DISULFIDE 65 BY SIMILARITY.
FT	DISULFID 58 BY SIMILARITY.
FT	DISULFID 205 BY SIMILARITY.
FT	CARBODY 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBODY 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBODY 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	UNSURE 2 OR D.
SQ	SEQUENCE 221 AA; 24.41 MW; 3035D7870EA743DB CRC64;
Query Match	75.2% Score 79; DB 1; Length 221;
Best Local Similarity	70.0%; Pred. No. 6.1e-06;
Matches	14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
CC	1 VPOSIDWRDGSAYTWSVNGQ 20
Db	3 LPDSIDWREKGAVPVNQG 22
RESULT 14	CATL RAT ID_CATL_RAT STANDARD; PRT; 334 AA.
AC	P07154; Q9QV07; 07-APR-1988 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP) (Cyclic protein-2) (CP-2).
GN	CTSL.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxId	10116;
RN	SEQUENCE FROM N.A.; TISSUE=Kidney; STRAIN=Nistar; PR00105; PAPAIN.
RP	SEQUENCE FROM N.A.; TISSUE=Kidney; STRAIN=Sprague-Dawley; MEDLINE=90092543; PubMed=2599113;
RP	MEDLINE=9030047; PubMed=3666143;
RA	Ishidoh K., Towatari T., Imajoh S., Kawasaki H., Kominami E., Katunuma N., Suzuki K.,
RA	"Gene structure and 5'-upstream sequence of rat cathepsin L.";
RT	"Molecular cloning and sequencing of cDNA for rat cathepsin L.";
RL	FEBS Lett. 223:69-73 (1987).
RN	[2]
RP	SEQUENCE OF 1-42 DAWLEY; STRAIN=Sprague-Dawley; MEDLINE=2125611; PubMed=11356678;
RP	Zablioukov S.D., Charron M., Deerbo J.N., Simukova N., Wright W.W.; "Male germ cells regulate transcription of the cathepsin L gene by rat Sertoli cells.";
RT	FEBS Lett. 259:71-74 (1989).
[3]	SEQUENCE OF 1-42 DAWLEY; STRAIN=Sprague-Dawley; MEDLINE=2125611; PubMed=11356678;
RP	Zablioukov S.D., Charron M., Deerbo J.N., Simukova N., Wright W.W.; "Male germ cells regulate transcription of the cathepsin L gene by rat Sertoli cells.";
RT	Endocrinology 142:2318-2327 (2001).
RN	[4]
RP	SEQUENCE OF 88-334 FROM N.A.; TISSUE=Sertoli Cells; MEDLINE=2166015; PubMed=1791830;
RA	Carlson-Lawrence M., Zablioukov S.D., Wright W.W.; "Cyclic protein-2, a secretory product of rat Sertoli cells, is the proenzyme form of cathepsin L.";
RT	Mol. Endocrinol. 5:1789-1798 (1991).
RN	[5]
SEQUENCE OF 18-37, FUNCTION, AND SUBCELLULAR LOCATION.	
RP	STRAIN=Sprague-Dawley; TISSUE=Sertoli cells;
RC	RA Boujrad N., Ogwembu S.O., Garnier M., Lee C.-H., Martin B.M., Papadopoulou V.
RT	"Identification of a stimulator of steroid hormone synthesis isolated from testis.";
RT	Science 268:1609-1612 (1995).
RN	[6]
SEQUENCE OF 18-28, AND TISSUE SPECIFICITY.	
RP	STRAIN=Sprague-Dawley; TISSUE=Epidermis;
RC	RA MEDLINE=2014186; PubMed=10699763;
RX	Kawada A., Hara K., Kominami E., Tezuka T., Takahashi M., Takahara H.; "Precursor of rat epidermal cathepsin L: purification and immunohistochemical localization.";
RT	RT immunohistochemical localization.";
RT	RT Dermatol. Sci. 21:36-45 (2000).
RN	[7]
SEQUENCE OF 114-288 AND 291-334.	
RP	TISSUE=Liver;
RC	RA MEDLINE=88296890; PubMed=3402618;
RX	Towatari T., Katunuma N.; "Amino acid sequence of rat liver cathepsin L.";
RT	RT FEBS Lett. 236:57-61 (1988).
CC	-!- FUNCTION: Important for the overall degradation of proteins in lysosomes. Procathepsin L is required for maximal stimulation of steroidogenesis by TIMP1.
CC	-!- CATALYTIC ACTIVITY: Specificity close to that of papain. As compared to cathepsin B, cathepsin L exhibits higher activity towards protein substrates, but has little activity on Z-Arg Arg-NHMe, and no peptidyl-dipeptidase activity.
CC	-!- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE BONDS.
CC	-!- SUBCELLULAR LOCATION: Procathepsin L is secreted. Mature cathepsin L is lysosomal.
CC	-!- TISSUE SPECIFICITY: Both mature cathepsin L and procathepsin L are found in the upper epidermis. The lower epidermis predominantly contains procathepsin L. In seminiferous tubules expression is greater at stages VI-VII than at stages IX-XII.
CC	-!- INDUCTION: Expression in Sertoli cells is repressed by germ cells.
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY Cl.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC	between the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
CC	CC EMBL; Y00679; CAA6891.1; -.
CC	CC EMBL; AF025416; AAB1616.1; -.
CC	CC PRINTS; S85154; AAB21516.1; -.
CC	CC PIR; S07098; KARTL.
CC	CC HSPP; O60911; 1PH0.
CC	CC MEROPS; CO1.032; -.
CC	CC InterPro; IPR000668; Peptidase_C1.
CC	CC InterPro; IPR00169; S9prot_acsite.
CC	CC PFM0112; Peptidase_C1; 1.
CC	CC PRINTS; PR00705; PAPAIN.
CC	CC PRODOM; P000158; Peptidase_C1; 1.
CC	CC PROSITE; PS00137; THIOL_PROTASE_CYS; 1.
CC	CC PROSITE; PS00639; THIOL_PROTASE_HIS; 1.
CC	CC PROSITE; PS00640; THIOL_PROTASE ASN; 1.
KW	Hydrolase; Thiol protease; Glycoprotein; Lysosome; Signal.
FT	SIGNAL 1 17 ACTIVATION PEPTIDE.
FT	PROPEP 18 113 CATEPSIN L, HEAVY CHAIN.
FT	CHAIN 114 288 CATEPSIN L, LIGHT CHAIN.
FT	CHAIN 291 334 CATEPSIN L, LIGHT CHAIN.
FT	ACT SITE 138 138 BY SIMILARITY.
FT	ACT SITE 276 276 BY SIMILARITY.
FT	ACT SITE 300 300 BY SIMILARITY.
FT	DISULFID 169 211 BY SIMILARITY.

Mon Nov 25 08:17:59 2002

us-09-674-738-2.rsp

Page 10 .

FT HELIX 201 210
FT TURN 211 211
FT STRAND 213 213
FT STRAND 215 215
FT TURN 216 216
FT HELIX 231 233
FT STRAND 238 238
FT STRAND 242 245
FT HELIX 251 260
FT STRAND 263 267
FT HELIX 272 276
FT STRAND 281 282
FT STRAND 292 299
FT STRAND 303 307
FT STRAND 310 310
FT TURN 312 313
FT TURN 315 315
FT STRAND 316 316
FT TURN 317 318
FT STRAND 319 323
FT HELIX 332 334
FT TURN 335 336
FT STRAND 340 343
SQ SEQUENCE 345 AA; 38922 MW; 82D9PPB35EDCA12EF CRC64;
Query Match 75.2%; Score 79; DB 1; Length 345;
Best Local Similarity 65.0%; Pred. No. 9.9e-06;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy :|:|||:|||:|||:|||:
Db 1 VPGSIDWRSQGAVTSVKHQG 20
Db 134 IPEVVDWRQKGAVTPVKHQG 153

Search completed: November 21, 2002, 16:09:02
Job time : 8.5 secs

Query Match 100.0%; Score 108; DB 1; Length 212;
 Best Local Similarity 100.0%; Pred. No. 2.5e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTsvknQ 20
 Db 2 VPOSIDWRYGAVTsvknQ 21

RESULT 2
 ANAN_ANACO STANDARD; PRT; 345 AA.
 AC P8084; O22293; 35, Created
 DT 01-NOV-1997 (Rel. 16-Oct-2001 (Rel. 40, Last sequence update)
 DE Ananain precursor (EC 3.4.22.31).
 GN ANI.
 OS Ananas comosus (Pineapple).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 Bromeliaceae; Ananas.
 NCBI_TaxID=4615;
 OX [1];
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Smooth Cayenne; TISSUE=stem;
 RA Lee K.L., Albee K.L., Bernasconi R.J., Edmunds T.;
 RT "Complete amino acid sequence of ananain and a comparison with stem
 bromelain and other plant cysteine proteases.",
 RL Biochem. J. 327:159 (1997).
 CC -|- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 for peptide bonds. Best reported small molecule substrate Br-Phe-
 Val-Arg-1-NHMe, but broader specificity than fruit bromelain.
 CC -|- ENZYME REGULATION: STRONGLY INHIBITED BY CHICKEN EGG-WHITE
 CC CYSTATIN AND THE ACTIVE-SITE-DIRECTED INHIBITOR TRANS-
 CC EPOYSUCCINYL-L-LEUCYAMIDO-(4-GUANIDINO) BUTANE (E-64).
 CC -|- TISSUE SPECIFICITY: STEM.
 CC -|- MASS SPECTROMETRY: MW=234.78; METHOD=Electrospray.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to licenses@isb-sib.ch.)

DR EMBL:AJ002477; CAA05487.1; -
 DR MEROPS: C01.026; -
 DR InterPro: IPR00668; Peptidase Cl.
 DR InterPro: IPR00169; SHprt_acSite.
 DR Pfam: PF00112; Peptidase_Cl_1.
 DR PRINTS: PRO00705; PAPAIN.
 DR Prodrom: PD000158; Peptidase_Cl_1.
 DR PROSITE: PS00139; THIOL_Protease_Cys_1.
 DR PROSITE: PS00639; THIOL_Protease_His_1.
 DR PROSITE: PS00640; THIOL_Protease ASN_1.
 KW Hydrolase; Thiol protease; Zymogen; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT PROPP 25 122 CHAIN 123 345 ANANAIN.
 FT ACT_SITE 147 147 BY SIMILARITY.
 FT ACT_SITE 279 279 BY SIMILARITY.
 FT ACT_SITE 144 144 BY SIMILARITY.

FT DISULFID 178 217 BY SIMILARITY.
 PT DISULFID 273 325 BY SIMILARITY.
 FT CONFLICT 291 324 S -> I (IN REF. 2).
 FT CONFLICT 324 324 L -> I (IN REF. 2).
 SQ SEQUENCE 345 AA; 38248 MW; FAF2999080174D87 CRC64;

Query Match 86.1%; Score 93; DB 1; Length 345;
 Best Local Similarity 94.7%; Pred. No. 1.1e-07; 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
 Qy 1 VPOSIDWRYGAVTsvknQ 19
 Db 123 VPOSIDWRYGAVTsvknQ 141

Query Match 86.1%; Score 93; DB 1; Length 345;
 Best Local Similarity 94.7%; Pred. No. 1.1e-07; 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTsvknQ 19
 PRT; 368 AA.
 ID RD19_ARATH RD19_ARATH STANDARD; PRT; 368 AA.
 AC P43796; 32, Created
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Cysteine Proteinase RD19A precursor (EC 3.4.22.-).
 GN RD19A OR ATG39090 OR F19H02.190
 OC Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702; TAXID=3702;

SEQUENCE FROM N.A.
 STRAIN=cv. Columbia,
 MEDLINE=9314966; PubMed=8305504;
 Kozumi M., Yamaguchi-Shinozaki K., Tsuji H., Shinozaki K.;
 "Structure and expression of two genes that encode distinct drought-
 inducible cysteine proteinases in Arabidopsis thaliana.";
 Gene 129:175-182 (1993).
 [2]

SEQUENCE FROM N.A.
 STRAIN=cv. Columbia,
 MEDLINE=9314966; PubMed=8305504;
 Mayer K.F.X., Schueler C., Wambutt R., Murphy G., Volckaert G.,
 Pohl T., Duestehoefl A., Strelkina W., Entian K.-D., Terry N.,
 Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
 Weichselgartner M., de Simonne V., Obermaier B., Mache R., Mueller M.,
 Kreis M., Deisen M., Peijnenburg P., Watson M., Schmidheini T.,
 Reichert B., Portereit D., Perez-Alonso M., Bourry M., Bancroft I.,
 Vos P., Hoheisel J., Zimmermann W., Wedler H., Robben J.,
 Laingham S.-A., McCullagh B., Bilham L., Robben J.,
 Van der Schueren J., Gromponez B., Chuang Y.-J., Vandembussche F.,
 Braken M., Weltjens I., Voege M., Baetelaars I., Hert R., Defoor E.,
 Weitzengraber T., Boche G., Ramsperger U., Hilbert H., Braun M.,
 Holzer B., Brandt A., Peters S., van Staveren M., Dirkse W.,
 Moerman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 Berneirer S., Hempel S., Peldpausch M., Lambirth S., Van den Daele H.,
 De Keyser A., Buyschaert C., Gielen J., Villaruel R., De Clercq R.,
 Van Montagu M., Rogers J., Bray Allen S., Quail M.,
 Clark L., Doggett J., Hall S., Kay M., Lennard N., McElroy K., Mayes R.,
 Pertett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloecher H., Schärfe M., Grimm M., Loehnert T.-R.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Aur S.,
 RA Gabel C., Fuchs M., Farman B., Grandérath K., Dauner D., Herzl A.,
 RA Neumann S., Arcuriou A., Vitale D., Ligouri R., Piravandi E.,
 RA Massenreiter O., Quigley F., Cibaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Cooke R., Berger C., Monfort A., Auborg S.,
 RA Chef d'or F., Gibbons T., Weber N., Vandembol M., Bérgues M.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Torres A.,
 RA Heijnen L., Schwart S., Scholler P., Heber S., Francis P., Bieleke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bayan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dediha N., Gnoj L., Schutz K., Abu-Threideh J.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA

Stonkeling T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Maricle E., Dante M., Peigin K., Hillier L., Nelson J., Spiehler J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C., Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher N., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen B., Maria M., Martienense R., McCombie W.R.; RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."; RL Nature 402:769-777(1999).
 CC -!- INDUCTION: BY HIGH SALT CONDITIONS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY CL.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AL05679; CAR38829; 1; -.
 DR EMBL; AL163594; CABR0572; 1; -.
 DR EMBL; D13042; BAA02373; 1; -.
 DR EMBL; P07711; ICJL.
 DR MEROPS; C01_002; -.
 DR INT-PRO; IPR00668; Peptidase_C1.
 DR InterPro; IPR000169; SHProt_acsite.
 DR PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00159; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00139; THIOL_Protease_His; 1.
 DR PROSITE; PS00630; THIOL_Protease_Asn; 1.
 DR PROSITE; PS00640; THIOL_Protease_Signal; 1.
 KW Hydrolase; Thiol_protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 134 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 135 368 CYSTEINE PROTEINASE RD19A.
 FT ACT_SITE 159 BY SIMILARITY.
 FT ACT_SITE 302 BY SIMILARITY.
 FT ACT_SITE 329 BY SIMILARITY.
 FT DISULFID 156 206 BY SIMILARITY.
 FT DISULFID 194 240 BY SIMILARITY.
 FT DISULFID 296 350 BY SIMILARITY.
 FT CARBOHYDROXY 253 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 368 AA; 40418 MW; 9D64CF35E807F519D CRC64;
 Query Match 72.3%; Score 78; DB 1; Length 368;
 Best Local Similarity 68.4%; Pred. No. 3.2e-05;
 Matches 13; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPOSIDWDRDGAVTSVKQ 19
 QY 1 :|||:|||||:||||| 19
 Db 135 LPEDFDWRDHGAVTPVKQ 153

RESULT 4
 PAP2_CAMP, STANDARD; PRT; 352 AA.
 ID PAP2_CARP, STANDARD; PRT; 352 AA.
 AC P14080;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymopain precursor (EC 3.4.22.6) (Papaya proteinase II) (PPII).
 OS Carica papaya (Papaya).
 OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucommias; II; Brassicales; Caricaceae; Carica.
 OC NCBI_TAXID=3649;

RESULT 5
 CATAL_PIG, STANDARD; PRT; 334 AA.
 ID CATAL_PIG, STANDARD; PRT; 334 AA.
 AC Q28944;

RN [1] SEQUENCE FROM N.N.A.
 RP TISSUE-leaf:
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Maricle E., Dante M., Peigin K., Hillier L., Nelson J., Spiehler J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C., Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher N., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen B., Maria M., Martienense R., McCombie W.R.; RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."; RL Nature 402:769-777(1999).
 CC Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases. [2].
 RN SEQUENCE OF 135-352.
 RP MEDLINE:9017930; PubMed:2106878;
 RX "The amino acid sequence of chymopapain from Carica papaya.";
 RA Biocchem. J. 266:75-81(1990).
 RN [3] SEQUENCE OF 135-352.
 RP MEDLINE:89302685; PubMed:2500950;
 RX Jacquet A., Kleinschmidt T., Schoneck A.G., Loosze Y., Braunitzer G.; RT "The thiol proteases from the latex of Carica papaya L. III. The primary structure of chymopapain.";
 RL Biol. Chem. Hoppe-Seyler 370:425-434(1989).
 RN [4] X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RP MEDLINE:9712865; PUBMED:897203;
 RX Maes D., Bouckaert J., Poortmans F., Wyns L., Loosze Y.; RT "Structure of chymopapain at 1.7-A resolution.";
 RL Biochemistry 35:1629-1629(1996).
 CC "Structure of chymopapain at 1.7-A resolution.";
 CC Maes D., Bouckaert J., Poortmans F., Wyns L., Loosze Y.; RT "Structure of chymopapain at 1.7-A resolution.";
 CC DR MEROPS; C01_002; -.
 DR InterPro; IPR00668; Peptidase_C1.
 DR PIR; S04222; S04222.
 DR PIR; S08285; S08285.
 DR PDB; 1YAL; 23-DEC-96.
 DR MEROPS; C01_002; -.
 DR InterPro; IPR00668; Peptidase_C1.
 DR PIR; S04222; S04222.
 DR PIR; S08285; S08285.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00139; THIOL_Protease_CYS; 1.
 DR PROSITE; PS00639; THIOL_Protease_His; 1.
 DR PROSITE; PS00640; THIOL_Protease_Asn; 1.
 DR PROSITE; PS00640; THIOL_Protease_Signal; 1.
 DR PROSITE; PS00640; THIOL_Protease_Signal; 3D-structure.
 KW Hydrolase; Thiol_protease; Zymogen; Signal; 3D-structure.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 134 ACTIVATION PEPTIDE.
 FT CHAIN 135 352 CHYMOPAPAIN.
 FT ACT_SITE 159 159 BY SIMILARITY.
 FT ACT_SITE 293 293 BY SIMILARITY.
 FT DISULFID 313 313 BY SIMILARITY.
 FT DISULFID 156 197 BY SIMILARITY.
 FT DISULFID 190 229 BY SIMILARITY.
 FT DISULFID 287 338 BY SIMILARITY.
 SQ SEQUENCE 352 AA; 39414 MW; 50E31EBFCFOA9F CRC64;
 Query Match 70.4%; Score 76; DB 1; Length 352;
 Best Local Similarity 83.3%; Pred. No. 6.4e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Cathepsin L precursor (EC 3.4.22.15).
 GN CTSL
 OS Sus scrofa (Pig).
 OC Mammalia; Eutheria; Chorata; Craniata; Vertebrate; Buteleostomi;
 OC NCBI_TaxID=9823;
 RN [1] -
 RP SEQUENCE FROM N.A.
 RC TISSUE=Epithelium; PubMed=7492581;
 RX MEDLINE=16059841;
 RA Okamura N., Tamba M., Uchiyama Y., Sugita Y., Dacheux F.,
 RA Syntin P.; Dacheux J.L.;
 RT "Direct evidence for the elevated synthesis and secretion of procathepsin L in the distal caput epididymis of boar.";
 RL Biochim. Biophys. Acta 1245:221-226 (1995).
 CC -!- FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN LYSOMES.
 CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain. As compared to cathepsin B, cathepsin L exhibits higher activity towards protein substrates, but has little activity on Z-Arg-Arg-NHMeC, and no peptidyl-dipeptidase activity.
 CC -!- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal. BELONGS TO PEPTIDASE FAMILY C1.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR BAA07917; BA007140.1; -.
 DR MEROPS; C01.032; -.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprtase_acsite.
 PFam; PF00112; Peptidase_C1; -.
 PRINTS; PR00705; PAPAIN_1.
 DR PD000158; Peptidase_C1; -.
 DR PROSITE; PS00139; THIOL_PROTEASE_HIS_1.
 DR PROSITE; PS00639; THIOL_PROTEASE ASN_1.
 KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 17 ACTIVATION PEPTIDE.
 FT PROPP 18 117 CATHESPIN_L_HEAVY_CHAIN.
 FT CHAIN 118 289 BY SIMILARITY.
 FT PROPEP 290 291 CATHESPIN_L_LIGHT_CHAIN.
 FT CHAIN 292 334 BY SIMILARITY.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT ACT_SITE 277 277 BY SIMILARITY.
 FT ACT_SITE 301 301 BY SIMILARITY.
 FT DISULFID 135 178 BY SIMILARITY.
 FT DISULFID 169 212 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 270 323 N-LINKED (GLCNAC -) (POTENTIAL).
 FT CARBOHYD 222 222 N-LINKED (GLCNAC -) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC -) (POTENTIAL).
 SQ 334 AA; 37178 MW; 51DBA79ACCF2CE53 CRC64;

Query Match 69.4%; Score 75; DB 1; Length 334;
 Best Local Similarity 68.4%; Pred. No. 8.9e-05;
 Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

CC 1 VPOSIDWRYGAVTSVKQ 19
 DB 114 VPKSVWDREKGTVTAKQ 132

RESULT 6
 CC3_CARCN STANDARD; PRT; 43 AA.
 ID CC3_CARCN
 AC P32566;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update) (Fragment).
 DE Cysteine proteinase III (EC 3.4.22.-) (CC-III)
 OS Carica candaemarcensis.
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euroids II; Brassicales; Caricaceae; Carica.
 OC euroids II; Brassicales; Caricaceae; Carica.
 RN [1] -
 RP Sequence; TISSUE=Latex;
 RX MEDLINE=94030669; PubMed=8216902;
 RA Walravens V., Jaziri M., van Beeumen J., Schnek A.G., Kleinschmidt T., Loosse Y.;
 RT Isolation and preliminary characterization of the cysteine-proteinases from the latex of Carica candaemarcensis Hook.;
 RL Biol. Chem. Hoppe-Seyler 374:501-506 (1993).
 CC -!- PTM: GLYCOSYLATED.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 DR HSSP; P14080; 1YAL.
 DR MEROPS; C01.020; -.
 DR IntePro; IPR000668; Peptidase_C1.
 DR InterPro; IPR00169; SHprtase_acsite.
 DR Pfam; PF00112; Peptidase_C1; -.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS_PARTIAL.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN_PARTIAL.
 DR HYDROLASE; PS00139; THIOL_Protease_CYS_1.
 KW Hydrolase; Thiol protease; Glycoprotein.
 FT ACT_SITE 25 25 BY SIMILARITY.
 FT NON_TER 43 43 BY SIMILARITY.
 SQ SEQUENCE 43 AA; 4636 MW; F4CD281886E291 CRC64;

Query Match 68.5%; Score 74; DB 1; Length 43;
 Best Local Similarity 77.8%; Pred. No. 1.5e-05;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CC 2 POSIDWRYGAVTSVKQ 19
 DB 2 PESIDWRYGAVTSVKQ 19

RESULT 7
 CC4_CARCN STANDARD; PRT; 43 AA.
 ID CC4_CARCN
 AC P32567;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DE Cysteine proteinase IV (EC 3.4.22.-) (CC-IV) (Fragment).
 OS Carica candaemarcensis.
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euroids II; Brassicales; Caricaceae; Carica.
 OC euroids II; Brassicales; Caricaceae; Carica.
 RN [1] -
 RP Sequence; TISSUE=Latex;
 RX MEDLINE=94030669; PubMed=8216902;
 RA Walravens V., Jaziri M., van Beeumen J., Schnek A.G., Kleinschmidt T., Loosse Y.;
 RT Isolation and preliminary characterization of the cysteine-proteinases from the latex of Carica candaemarcensis Hook.;
 RL Biol. Chem. Hoppe-Seyler 374:501-506 (1993).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 DR HSSP; C01.020; 1YAL.
 DR IntePro; IPR000668; Peptidase_C1.

DR InterPro; IPR000169; Shprt_acete.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00539; THIOL_PROTEASE_HIS; PARTIAL.
 DR PROSITE; PS00590; THIOL_PROTEASE ASN; PARTIAL.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR Hydrolase; Thiol protease.
 FT ACT SITE 25 25 BY SIMILARITY.
 FT NON_TER 43 43 .
 SQ 43 AA; 4685 MW; FAD5945A9386E291 CRC64;
 Query Match 68.5%; Score 74; DB 1; Length 43;
 Best Local Similarity 77.8%; Pred. No. 1.5e-05;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OX 2 POSIDWDRDGAVGAVSKQ 19
 DB 2 PESIDWDRKGAVGAVPKQ 19 .
 RESULT 8
 CATH_MOUSE STANDARD; PRT; 334 AA.
 ID CATH_MOUSE .
 AC P06737; 01-JAN-1998 (Rel. 06, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DR 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP).
 GN CTSL.
 OS Mus musculus (Mouse).
 OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelostomi.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10930;
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=83076349; PubMed=6891318;
 RA Tzien R.R., Gil S., Gottstein M.M.;
 RR "Sequence and expression of the cDNA for MEP (major excreted
 RR protein), a transformation-regulated secreted cathepsin.";
 RL Biochem. J. 246:731-735(1987).
 RN [2] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=B8213715; PubMed=2835398;
 RA Joseph L.J., Chang L.C., Stamenkovich D., Sukharev V.P.;
 RR "Complete nucleotide and deduced amino acid sequences of human and
 RR murine procathepsin L. An abundant transcript induced by
 RR transformation of fibroblasts.";
 RL Clin. Invest. 81:1621-1629(1988).
 RN [3] SEQUENCE FROM N.A.,
 RR MEDLINE=87033683; PubMed=3533924;
 RA Portnoy D.A., Erickson A.H., Kochan J., Ravetch J.V., Unkeless J.C.;
 RR "Cloning and characterization of a mouse cysteine proteinase.";
 RL J. Biol. Chem. 261:14697-14703(1986).
 RN [4] SEQUENCE FROM N.A., AND CARBOHYDRATE-LINKAGE SITE.
 RC TISSUE=Liver;
 RK MEDLINE=9111261; PubMed=2275556;
 RA Stearns N.A., Dong J., Pan J.X., Brenner D.A., Sahagian G.G.;
 RR "Comparison of cathepsin L synthesized by normal and transformed
 RR cells at the gene, message, protein, and oligosaccharide levels.";
 RL Arch. Biochem. Biophys. 283:447-457(1990).
 RN [5] SEQUENCE OF 89-300 FROM N.A.
 RC STRAIN=BALB/C;
 RK MEDLINE=86271744; PubMed=3755373;
 RA Denhardt D.T., Hamilton R.T., Parfett C.L.J., Edwards D.R.,
 RA Pierre R.S., Waterhouse P., Nilson-Hamilton M.;
 RR "Close relationship of the major excreted protein of transformed
 RR murine fibroblasts to thiol-dependent cathepsins.";
 RL Cancer Res. 46:4590-4593(1986).
 CC -!- FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN
 LYOSOMES.

CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain. As compared to cathepsin B, cathepsin L exhibits higher activity towards protein substrates, but has little activity on Z-Arg-Arg-NHMe, and no peptidyl-dipeptidase activity.
 CC -!- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X06086; CR29470.1; --.
 CC DR EMBL; J02583; AAA37445.1; --.
 CC DR EMBL; M2045; AAA39984.1; --.
 CC DR EMBL; X04392; CR22780.1; --.
 CC DR PIR; S01177; KHSNL.
 CC DR PIR; S13890; S13890.
 CC DR HSSP; P07711; 1CIL.
 CC DR MEROPS; C01.032; --.
 CC DR MGD; M88564; Ctsl.
 CC DR InterPro; IPR000668; Peptidase_C1.
 CC DR InterPro; IPR000169; Shprt_acete.
 CC DR Pfam; PF00112; Peptidase_C1; 1.
 CC DR PRIMIS; PR00705; PAPRIN.
 CC DR PRODOM; PD000158; Peptidase_C1; 1.
 CC DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 CC DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 CC DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 CC KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 CC FT SIGNAL; 1 17 POTENTIAL.
 CC FT PROPEP; 18 113 ACTIVATION PEPTIDE.
 CC FT CHAIN; 114 288 CATHESPIN_L, HEAVY CHAIN.
 CC FT CHAIN; 291 334 CATHESPIN_L, LIGHT CHAIN.
 CC FT ACT_SITE; 138 288 BY SIMILARITY.
 CC FT ACT_SITE; 276 276 BY SIMILARITY.
 CC FT ACT_SITE; 300 300 BY SIMILARITY.
 CC FT DISULFO; 135 178 BY SIMILARITY.
 CC FT DISULFO; 169 211 BY SIMILARITY.
 CC FT DISULFO; 269 322 INTERCHAIN (BY SIMILARITY).
 CC FT CARRYHO; 221 221 N-LINKED (GLCNAC. .).
 CC FT CONFLICT; 58 58 M -> I (IN REF. 2).
 CC FT CONFLICT; 177 177 G -> R (IN REF. 3).
 CC SQ SEQUENCE 334 AA; 37547 MW; FE674704307A998 CRC64;
 Query Match 68.5%; Score 74; DB 1; Length 334;
 Best Local Similarity 63.2%; Pred. No. 0.00013; Matches 12; Mismatches 4; Indels 0; Gaps 0;
 OX 1 VPOSIDWDRDGAVGAVSKQ 19
 DB 114 IPKSVWDREKGCVTVKQ 132 .
 RESULT 9
 ID A494_ARATH STANDARD; PRT; 361 AA.
 AC P43395; Q9S7M5; DT 01-NOV-1995 (Rel. 32, Created)
 AC P43395; Q9S7M5; DT 16-OCT-2001 (Rel. 40, Last sequence update)
 AC P43395; Q9S7M5; DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable cysteine proteinase A494 precursor (EC 3.4.22.-).
 GN At2021430 OR F3K31_19
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Bokaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;
 OC Spermatophytina; Magnoliophytina; euicocyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:06:44 ; Search time 25 Seconds
(without alignments) 164.838 Million cell updates/sec

Title: US-09-674-738-2

Perfect score: 105 VPOSIDWRDGSAGTYSVKNQG 20

Sequence:

Scoring table: BIOSUM62

GapOp 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBLU 21:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_muc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_rabbit:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriophage:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	105	100.0	20	10 Q9S7J0	Q9S7J0 ananas como
2	94	89.5	20	10 Q9SBM1	Q9SBM1 ananas como
3	93	88.6	20	10 Q9S9E5	Q9S9E5 ananas como
4	93	88.6	356	10 081084	081084 ananas como
5	93	88.5	356	10 023799	023799 ananas como
6	88	83.8	454	10 04922	04922 pseudotsuga
7	85	81.0	334	5 Q9U554	Q9U554 tetra aura
8	84	80.0	226	10 Q9SMH9	Q9SMH9 carica papa
9	84	80.0	227	10 Q9SMF8	Q9SMF8 carica papa
10	84	80.0	324	10 023801	023801 ananas como
11	84	80.0	337	10 Q8S334	Q8S334 lycopepsico
12	84	80.0	351	10 023791	023791 ananas como
13	84	80.0	352	10 Q9SMI1	Q9SMI1 carica papa
14	84	80.0	361	10 Q9SM10	Q9SM10 carica papa
15	83	79.0	357	10 081085	081085 ananas como
16	83	79.0	367	10 040261	040261 mesembryant

ALIGNMENTS

RESULT 1

ID	Q9S7J0	PRELIMINARY;	PRT;	20 AA.
AC	Q9S7J0;	Created)		
DT	01-MAY-2000 (TREMBLU 13,	Last sequence update)		
DT	01-MAY-2000 (TREMBLU 19,	Last annotation update)		
DE	Bromelain (Fragment).			
OS	Ananas comosus (Pineapple).			
OC	Eukaryote; Viridiplantae; Streptophyta; Embryophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;			
OC	Bromeliaceae; Ananas.			
OX	NCBI_TaxID=4615;			
RN	[1] SEQUENCE FROM N.A.			
RX	MEDLINE=95298192; PubMed=7779362;			
RA	Harrach T., Eckert K., Schulze-Forster K., Nuck R., Grunow D.,			
RA	Maurer H.R.,			
RT	"Isolation and partial characterization of basic proteinases from stem			
RT	bromelain;"			
J	Protein Chem. 14:41-52(1995).			
RN	[2] SEQUENCE.			
RP	MEDLINE=94330946; PubMed=8053898;			
RX	Napper A.D., Bennett S.P., Borowski M., Holdridge M.B., Leonard M.J.,			
RA	Rogers E.B., Duan Y., Lausen R.A., Reinhold B., Shames S.L.,			
RA	Bichem. J. 301:727-735(1994).			
RL	InterPro: IPR00668: Peptidase_C1.			
DR	pfam: PF00112; Peptidase_C1; 1			
SEQUENCE	20 AA; 214 MW; 83A842BC812C3EBC CRC64;			
SQ	Query Match 100.0%; Score 105; DB 10; Length 20; Best Local Similarity 100.0%; Pred. No. 1_3e-10; Mismatches 0; Indels 0; Gaps 0;			
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1 VPOSIDWRDGSAGTYSVKNQG 20			
Db	1 VPOSIDWRDGSAGTYSVKNQG 20			

FT CHAIN 123 338 PBSB
 SQ SEQUENCE 356 AA; 39560 MW; 16C2DCCB47238928 CRC64;
 Query Match 88.6%; Score 93; DB 10; Length 356;
 Best Local Similarity 94.7%; Pred. No. 3.4e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPOSIDWRSQAVTSVKNG 19
 DB 124 VPSIDWRSQAVTSVKNG 142

RESULT 6

ID Q40922 PRELIMINARY; PRT; 454 AA.
 AC Q40922; Q40919;
 DT 01-NOV-1996 (TREMBrel. 01, Created)
 DT 01-MAY-1996 (TREMBrel. 01, Last sequence update)
 DR FSEUDOTZAIN (Cysteine protease).
 GN PM3CCYSP.
 OS Pseudotsuga menziesii (Douglas fir).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pseudotsuga.
 OX NCBI_TaxID=3357;
 RN [1]
 RR SEQUENCE FROM N.A.
 RC STRAIN=[MIRB] FRANCO;
 RX MEDLINE:96269408; PubMed:8682307;
 RA Tranbarcer T.J.; Misra S.;
 RT "Structure and expression of a developmentally regulated cDNA encoding
 a cysteine protease (pseudotzain) from Douglas fir.";
 RL Gene 172:221-226(1995).
 RN [2]
 RR SEQUENCE FROM N.A.
 RC STRAIN=[MIRB] FRANCO;
 RA Tranbarcer T.J.; Misra S.;
 RL Physiol. Plantarum 0:0-0(0).
 RN [3]
 RR SEQUENCE OF 297-347 FROM N.A.
 RC TISSUE=WHOLE SEEDED;
 RA Tranbarcer T.J.; Misra S.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; JLM192; AAC9455.1; -;
 DR EMBL; 249765; CAB98935.1; -;
 DR HSSP; P07711; 1CUL.
 MEROPS: C01-029; -;
 DR InterPro; IPR000118; Granulin.
 DR InterPro; IPR000169; SHPr0t_acsite.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000112; Peptidase_C1.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR SMART; SM0077; GRAN_1.
 DR PROSITE; PS00118; PA2_HIS_UNKOWN_1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 DR Hydrolase; Protease; Thiol_protease.
 SQ SEQUENCE 454 AA; 48CI93FD66CA023F CRC64;
 Query Match 83.8%; Score 88; DB 10; Length 454;
 Best Local Similarity 75.0%; Pred. No. 3.1e-06;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPOSIDWRSQAVTSVKNG 20
 DB 132 VPSIDWRSQAVTSVKNG 151

RESULT 7

FT CHAIN 1 PRELIMINARY; PRT; 334 AA.
 SQ SEQUENCE 1 AA; 1 MW; 1 PRELIMINARY;
 Query Match 80.0%; Score 84; DB 10; Length 226;
 DR OQU554 PRELIMINARY; PRT; 334 AA.
 AC OQU554;
 DT 01-MAY-2000 (TREMBrel. 13, Created)
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)
 DR Silicatein beta.
 OS Tethya aurantia.
 OC Hadromerida; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
 OC Nobi; TaxID=34494;
 RN [1]
 RR SEQUENCE FROM N.A.
 RA Shimizu S; Lawrence C.; Morse D.E.;
 RT "Cloning of silicatein beta.",
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF08670; AF22819.1; -;
 DR BSSP; P07711; 1CUL.
 DR InterPro; IPR00668; Peptidase_C1.
 DR InterPro; IPR000169; SHPr0t_acsite.
 DR PFAM; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS0039; THIOL_PROTEASE HIS; UNKNOWN 1.
 SQ SEQUENCE 334 AA; 36787 MW; 4E6754BB23799065 CRC64;
 Query Match 81.0%; Score 85; DB 5; Length 334;
 Best Local Similarity 78.9%; Pred. No. 6.9e-06;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 DR 119 VPSIDWRSQAVTSVKNG 137

RESULT 8

ID O9SMH9 PRELIMINARY; PRT; 226 AA.
 AC O9SMH9;
 DT 01-MAY-2000 (TREMBrel. 13, Created)
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)
 DR C1-MAR-2002 (TREMBrel. 20, Last annotation update)
 DE Chymopapain isoform IV (EC 3.4.22.6) (Fragment).
 GN CHYMOIV.
 OC Carica papaya (Papaya).
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosidis II; Brassicales; Caricaceae; Carica.
 OX NCBI_TaxID=3649;
 RR SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Taylor M.A., Al-Sheikh M., Revell D.F., Summer I.G., Connerton I.F.;
 RT "cDNA cloning and expression of Carica papaya prochymopapain isoforms
 in Escherichia coli",
 RL Plant Sci. 145: 41-47(1999).
 DR EMBL; AJ131997; CAB38316.1; -.
 DR HSSP; P14080; LYAL.
 MEROPS: C01-002; -;
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHPr0t_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1..
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1..
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1..
 DR Hydrolase; Protease; Thiol_protease.
 KW Hydrolase; Thiol_protease.
 FT NON_TER 1
 SQ SEQUENCE 226 AA; 24597 MW; 997A1B6B0CE49C35 CRC64;

Query Match 80.0%; Score 84; DB 10; Length 226;

Best Local Similarity 84.2%; Prod. No. 6.5e-06; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDRDSGAVTSPVKNQ 20
Db 1 POSIDRAKAGAVTPVKNQ 19

RESULT 9
Q9SMH8 PRELIMINARY; PRT; 227 AA.
ID Q9SMH8;
AC "CNA cloning and expression of Carica papaya prochymopapain isoforms in Escherichia coli";
DT 01-MAY-2000 (TREMBLrel. 1.3, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 2.0, Last annotation update)
DE Chymopapain isoform V (EC 3.4.22.6) (Fragment).
GN CHMOV.

OS Carica papaya (Papaya).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae;
euroids II; Brassicales; caricaeae; Caricae; Carica.
NCBI_TaxID:3649;
RN [1].

PP SEQUENCE FROM N.A.
RC Taylor M.A., A. Sheikh M., Revell D.F., Sumner I.G., Connerton I.F.;
RA RT "CNA cloning and expression of Carica papaya prochymopapain isoforms in Escherichia coli";
RT Plant Sci. 145:41-47 (1999);
RL AJ131198; CAB38317; -.
DR HSSP; P14080; YALV.
DR MEROPS; C01_0024;
DR InterPro; IPR00668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; -.
DR PRINTS; PRO0705; PAPAIN_.
DR PRODOM; PD00158; Peptidase_C1; 1.
DR PROSITE; PS00139; THIOL_Protease ASN; 1.
DR PROSITE; PS00639; THIOL_Protease_HIS; 1.
DR Hydrolase; Thiol protease.
FT NON_TER 1 1 MN; 1CC7C9D27AFAAAB0 CRC64;
SEQUENCE 227 AA; 24721 MN; 1CC7C9D27AFAAAB0 CRC64;

Query Match 80.0%; Score 94; DB 10; Length 227;
Best Local Similarity 84.2%; Prod. No. 6.5e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDRDSGAVTSPVKNQ 20
Db 2 POSIDRAKAGAVTPVKNQ 20

RESULT 10
O23801 PRELIMINARY; PRT; 324 AA.
ID O23801;
AC "CNA cloning and expression of Carica papaya prochymopapain isoforms in Escherichia coli";
DT 01-JAN-1998 (TREMBLrel. 0.5, Created);
DT 01-MAR-2002 (TREMBLrel. 0.5, Last sequence update)
DE FB1035 (Fragment).
OS Ananas comosus (Pineapple).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
OC Bromeliaceae; Ananas.
OX NCBI_TaxID:4615;
RN [1].

RP SEQUENCE FROM N.A.
RC STNAPPCV_N67-10; TISSUE=FRUIT;
RA Muta E., Okamoto Y., Ota S.;
RT "Cloning and sequencing of cysteine proteinases in Ananas comosus.";
RL Submitted (OCT-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL; D38534; BRA22546; -.

RESULT 11
Q8S34 PRELIMINARY; PRT; 337 AA.
ID Q8S34;
AC "A Tomato Cysteine Protease Required for Cf-2-Dependent Disease Resistance and Suppression of Autonecrosis.";
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Cysteine protease.
GN Lyco pericon pennellii (Tomato).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
NCBI_TaxID=28522;
RN [1].

RP SEQUENCE FROM N.A.
RX PubMed=1197458 A.
RA Kriger J.J., Thomas C.M., Golstein C., Dixon M.S., Smoker M., Tang S., Muider L., Jones J.D.;
RT "A Tomato Cysteine Protease Required for Cf-2-Dependent Disease Resistance and Suppression of Autonecrosis.";
RL Science 296:744-747 (2002);
DR EMBL; AF493233; AAC192091; -.
KW Protease.
SQ SEQUENCE 337 AA; 37389 MN; CE76884456564FBE CRC64;

Query Match 80.0%; Score 84; DB 10; Length 337;
Best Local Similarity 70.0%; Prod. No. 1e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDRDSGAVTSPVKNQ 20
Db 96 VPOSIDRDSGAVTSPVKNQ 114

RESULT 12
O23791 PRELIMINARY; PRT; 351 AA.
ID O23791;
AC "CNA cloning and expression of Carica papaya prochymopapain isoforms in Escherichia coli";
DT 01-JAN-1998 (TREMBLrel. 0.5, Created);
DT 01-MAR-1998 (TREMBLrel. 0.5, Last sequence update)
DE Bromelain precursor.
OS Ananas comosus (Pineapple).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
OC Bromeliaceae; Ananas.
OX NCBI_TaxID=4615;
RN [1].

RP SEQUENCE FROM N.A.
 RC STRAIN=CCV_N67-10; TISSUE=FRUIT;
 RA Muta E., Aramaki H., Takata Y., Kono A., Okamoto Y., Ota S.;
 RT "Cloning and sequencing of fruit bromelain."
 RL Submitted (JAN1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D14059; BA21849.1; -.
 DR HSSP; P07711; IYAL.
 DR MEROPS; C01-028; -.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000669; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PRO0705; PAPAIN.
 DR Prodrom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; UNKNOWN_1.
 KW Hydrolase; Signal; Thiol_protease.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 122 339 BROMELAIN.
 SQ SEQUENCE 351 AA; 39055 MW; 33781C55144242B0 CRC64;

Query Match 80.0%; Score 84; DB 10; Length 351;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDNRDGSATVSKNQ 19
 ||||| ||||| ||||| ||||| ||||| 141

RESULT 13

ID Q9SM11 PRELIMINARY; PRT; 352 AA.
 AC Q9SM11;
 DT 01-MAY-2000 (TREMBrel. 13, Created)
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)
 DE Chymopapain isoform III precursor (EC 3.4.22.6).
 GN CHM011.
 OS Carica papaya (Papaya).
 RC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 RT "CONA cloning and expression of Carica papaya prochymopapain isoforms
 in Escherichia coli.";
 RL Plant Sci. 145:41-47(1999).
 DR PROSITE; C01-002; -.
 DR InterPro; IPR000668; Peptidase_C1.
 DR PROSITE; PS00639; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PRO0705; PAPAIN.
 DR Prodrom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; 1.
 DR HSSP; P14080; IYAL.
 DR MEROPS; C01-028; -.
 DR InterPro; IPR000669; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PRO0705; PAPAIN.
 DR Prodrom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; 1.
 KW Hydrolase; Signal; Thiol_protease.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 135 361 CHYMPARIN III.
 SQ SEQUENCE 361 AA; 40478 MW; EBB2A6AE6D5CDC CRC64;

Query Match 80.0%; Score 84; DB 10; Length 361;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDWRDGSATVSKNQG 20
 ||||| ||||| ||||| ||||| ||||| 154

RESULT 14

ID Q9SM10 PRELIMINARY; PRT; 361 AA.
 AC Q9SM10;
 DT 01-MAY-2000 (TREMBrel. 13, Created)
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)
 DE Chymopapain isoform III precursor (EC 3.4.22.6).
 GN CHM011.
 OS Carica papaya (Papaya).
 RC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 RT "CONA cloning and expression of Carica papaya prochymopapain isoforms
 in Escherichia coli.";
 RL Plant Sci. 145:41-47(1999).
 DR PROSITE; C01-002; -.
 DR InterPro; IPR000668; Peptidase_C1.
 DR PROSITE; PS00639; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE HIS; 1.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PRO0705; PAPAIN.
 DR Prodrom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; 1.
 DR HSSP; P14080; IYAL.
 DR MEROPS; C01-028; -.
 DR InterPro; IPR000669; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PRO0705; PAPAIN.
 DR Prodrom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; 1.
 KW Hydrolase; Signal; Thiol_protease.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 135 352 CHYMPARIN II.

SQ SEQUENCE 352 AA; 39462 MW; A1EB21FABCFOB235 CRC64;

Query Match 80.0%; Score 84; DB 10; Length 352;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDWRDGSATVSKNQG 20
 ||||| ||||| ||||| ||||| 154

DR InterPro; IPR00169; SHprot_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PRO0705; PAPAIN
DR PRODom; PD000156; Peptidase_C1; 1.
DR PROSITE; PS000644; THIOL_PROTEASE ASN; 1.
DR PROSITE; PS000639; THIOL_PROTEASE HIS; UNKNOWN_1.
KW Signal_1.
PT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 357 CYSTINE PROTEINASE PRECURSOR, AN11.
SQ SEQUENCE 357 AA; 39606 MW; DB393DD035B47C31 CRC64;

Query Match 79.0%; Score 83; DB 10; Length 357;
Best Local Similarity 88.9%; Pred. No. 1.e-0;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VPQSIDWRDSSAVTsvkn 18
|||: |||: |||:
Db 123 VPQSIDMRNYGAVTsvkn 140

Search completed: November 21, 2002, 16:10:00
Job time : 26 secs

Copyright (c) 1993 - 2002 Compugen Ltd.
GenCore version 5.1.3

OM protein - protein search, using sw model
Run on: November 21, 2002, 16:06:44 ; Search time 25 Seconds
Title: US-09-674-738-1
Perfect score: 108
Sequence: I VPQSIDWRDYGAIVTSVKNQN 20
(without alignments)
164.838 Million cell updates/sec

Scoring table: BLOSUM62
Gpop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21: *
1: sp_archea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp Rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteria: *
17: sp_archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	108	100	0	Q89655 ananas comm
2	108	100	0	Q89655 ananas comm
3	108	100	0	Q89655 ananas comm
4	99	91	7	Q89655 ananas comm
5	99	91	7	Q89655 ananas comm
6	97	89	8	Q89655 ananas comm
7	93	86	1	Q89655 ananas comm
8	92	85	2	Q89655 ananas comm
9	90	83	3	Q89655 ananas comm
10	90	83	3	Q89655 ananas comm
11	80	74	1	Q95VA7 fasciola gi
12	80	74	1	Q95VA7 fasciola gi
13	79	73	1	Q95VA7 fasciola gi
14	79	73	1	Q95VA7 fasciola gi
15	77	71	3	Q95VA7 fasciola gi
16	77	71	3	Q95VA7 fasciola gi

17	77	71.3	337	5	O9SV59	Q9yv59 carica radic
18	76	70.4	226	10	O9SMH8	Q9smh8 carica papaya
19	76	70.4	227	10	O9SMH8	Q9smh8 carica papaya
20	76	70.4	352	10	O9SM10	Q9sm10 carica papaya
21	76	70.4	361	10	O9SM10	Q9sm10 carica papaya
22	76	70.4	363	10	O9XGH8	Q9xgh8 nicotiana t
23	76	70.4	363	10	O9XGT9	Q9xgt9 nicotiana t
24	76	70.4	366	10	O9AUC5	Q9auc5 ipomoea batatas
25	76	70.4	366	10	O9MTD5	Q9mtd5 ipomoea batatas
26	76	70.4	369	10	O9MTD5	Q9mtd5 ipomoea batatas
27	75	69.4	189	5	Q96ET7	Q96et7 spironometra
28	75	69.4	208	10	O94BS1	Q94bs1 daucus carota
29	75	69.4	306	5	O9R4G4	Q9r4g4 fasciola hepatica
30	75	69.4	326	5	O9NGW2	Q9ngw2 fasciola gigantica
31	75	69.4	326	5	O24944	Q24944 fasciola hepatica
32	75	69.4	336	5	O9NF1	Q9nf1 sbironometra
33	75	69.4	336	5	O02586	Q02586 spironometra
34	75	69.4	360	10	O9FT3	Q9ft3 oya sativa
35	75	69.4	360	10	O945R8	Q945r8 sandersonia
36	75	69.4	365	10	O945B0	Q945b0 nicotiana tabacum
37	75	69.4	367	10	O9D261	Q9d261 mesembryanthemum heteroderera
38	75	69.4	374	5	O18455	Q18455 heteroderera
39	75	69.4	381	5	O9GNG6	Q9gng6 leishmania
40	75	69.4	394	5	O95WR6	Q95wr6 leishmania
41	75	69.4	443	5	P91628	P91628 leishmania
42	75	69.4	443	5	O95WR7	Q95wr7 leishmania
43	75	69.4	443	5	O18698	O18698 leishmania
44	75	69.4	443	5	P92185	P92185 leishmania
45	75	69.4	498	5	O16454	O16454 caenorhabditis elegans
						ALIGNMENTS
RESULT 1						
O9S955		PRELIMINARY;	PRT;	20 AA.		
ID	O9S965					
AC	O9S965;	(TREMBLrel. 1.3, Created)				
DT	01-MAY-2000	(TREMBLrel. 1.3, Last sequence update)				
DT	01-DEC-2001	(TREMBLrel. 1.9, Last annotation update)				
DE	Bromelain (Fragment).					
OS	Ananas comosus (Pineapple).					
OC	Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; Liliopsida; Commelinidae incertae sedis; Bromeliidae; Ananas.					
OX	NCBI_TAXID=4615;					
RN	[1]	SEQUENCE FROM N.A.				
RP		MDLINE#93298112; PubMed#7779262;				
RX		RA: Harrach T.; Eckert K.; Schulze-Forster K.; Nuck R.; Grunow D.,				
RA		Maurer H.R.;				
RT		'Isolation and partial characterization of basic proteinases from stem bromelain.'				
RT		J. Protein Chem. 14:41:52(1995).				
DR		Interrito; IPR00668; Peptidase_C1.				
DR		pflam; PF00112; Peptidase_C1; 1.				
SBQUENCE	2277 MW;	8F984AE812C3BBC CRC64;				
Query Match	100.0%	Score 108; DB 10; Length 20;				
Best Local Similarity	100.0%	Pred. No. 7, 1e-11;				
Matches 20;	Conservative 0;	Mismatches 0;				
OY	1	VPOSITIONDGVAVSKVON 20				
Db	1	VPOSITIONDGVAVSKVON 20				

DT 01-NOV-1998 (TREMBLrel. 08; Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20; Last annotation update)
 DE Cysteine proteinase, Ananas precursor.
 GN Ananas comosus (Pineapple).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromeliaceae; Ananas.
 OC NCBITaxonID=4615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SMOOTH CAYENNE; TISSUE=STEM;
 RA Robertson C.E., Goodenough P.W.;
 RT "Cloning and expression of ananain gene from pineapple.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AU009829; CAA08860.1; -.
 DR HSSP: P07711; 1CJL.
 DR MEROPS: C01_005; -.
 DR InterPro: IPR000668; Peptidase_C1.
 DR IntervPro: IPR000169; SHpot_acsite.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRMTS: PR00705; PAPAIN.
 DR PRODOM: PD000158; Peptidase_C1; 1.
 DR PROSITE: PS000640; THIOL_Protease ASN; 1.
 DR PROSITE: PS000139; THIOL_Protease HIS; 1.
 DR PROSITE: PS000639; THIOL_Protease HIS; UNKNOWN_1.
 DR Hydrolase; Signal; Thiol_protease.
 KW SIGNAL
 FT CYSTEINE PROTEINASE PRECURSOR, AN8.
 FT 26
 SQ 356 AA; 39528 MW; 97D705C62EC01DB8 CRC64;

Query Match Score 108; DB 10; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy ▶ 1 VPOSIDWRDYGAVTSYKNON 20
 Db 124 VPOSIDWRDYGAVTSYKNON 143

RESULT 3
 ID Q23799; PRELIMINARY; PRT; 356 AA.
 AC O23799;
 DT 01-JAN-1998 (TREMBLrel. 05; Created)
 DT 01-MAR-2002 (TREMBLrel. 20; Last annotation update)
 DE FBSS precursor.
 OS Ananas comosus (Pineapple).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromeliaceae; Ananas.
 OC NCBITaxonID=4615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SMOOTH CAYENNE; TISSUE=FRUIT;
 RA Mutu E., Okamoto Y., Ota S.;
 RT "Cloning and sequencing of cysteine proteinases in Ananas comosus.";
 RL Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: D38534; BA22546..1; -.
 DR HSSP: P07711; 1CJL.
 DR MEROPS: C01_028; -.
 DR InterPro: IPR000169; Peptidase_C1.
 DR PROSITE: PS000139; THIOL_Protease_CYS; 1.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR ProDom: PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_Protease ASN; 1.
 DR PROSITE; PS000639; THIOL_Protease HIS; UNKNOWN_1.
 DR Hydrolase; Thiol_protease.
 FT NON_TER
 SQ 95 AA; 36038 MW; 057462C59424339F CRC64;

Query Match Score 99; DB 10; Length 344;
 Best Local Similarity 90.0%; Pred. No. 5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy ▶ 1 VPOSIDWRDYGAVTSYKNON 20
 Db 96 VPOSIDWRDYGAVNEYKNON 115

RESULT 5
 ID Q23791; PRELIMINARY; PRT; 351 AA.
 AC O23791;
 DT 01-JAN-1998 (TREMBLrel. 05; Created)
 DT 01-MAR-2002 (TREMBLrel. 20; Last annotation update)
 DE Bromelain precursor.
 OS Ananas comosus (Pineapple).
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromeliaceae; Ananas.
 OC NCBITaxonID=4615;
 RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=CV.N67-10; TISSUE=FRUIT;

RN [1] SEQUENCE FROM N.A.
 RA RP
 RT SEQUENCE FROM N.A.
 RL MEDLINE=95299192; PubMed=779262;
 RL Submitted (JAN 1993) to the EMBL/GenBank/DDBJ databases.
 EMBL; DR
 DR HESP; DR07711; ICUL.
 DR MEROPS; C01_028; --.
 DR InterPro; IPR00668; Peptidase_C1.
 DR InterPro; IPR000169; SHProt_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_Protease_His.
 DR PROSITE; PS00159; THIOL_Protease_Cys; 1.
 KW Hydrolase; Signal; Thiol_protease.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 122 339 BROMELAIN.
 SQ SEQUENCE 351 AA; 39055 MW; 33781C55144242B0 CRC64;
 Query Match 91.7%; Score 99; DB 10; Length 351;
 Best Local Similarity 90.0%; Pred. No. 5.4e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OQ 1 VPOSIDWRYGAVTSVKNO 20
 Db 123 VPQSIDWRYGAVTSVKNO 142

RESULT 6
 Q9SM1 PRELIMINARY; PRT; 20 AA.
 ID Q98M1
 AC 001085; PRELIMINARY; PRT; 20 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE COMOSAN (Fragment).
 OC Ananas comosus (Pineapple).
 OC Buxaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromelaceae; Ananas.
 OX NCBI_TAXID=4615;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94310946; PubMed=8033898;
 RA Narper A.D., Bennett S.P., Borowski M., Holdridge M.B., Leonard M.J.,
 RA Rogers E.B., Duan Y., Laursen A., Reinhold B., Shames S.L.;
 RL Biochem. J. 301:727-735(1994).
 DR InterPro; IPR000169; Peptidase_C1.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR SEQUENCE 20 AA; 2219 MW; 83A84A16232C3EBC CRC64;

Query Match 89.8%; Score 97; DB 10; Length 20;
 Best Local Similarity 94.7%; Pred. No. 4.6e-09;
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OQ 1 VPOSIDWRYGAVTSVKNO 19
 Db 1 VPQSIDWRYGAVTSVKNO 19

RESULT 7
 Q9SM10 PRELIMINARY; PRT; 20 AA.
 ID Q9STJ0
 AC 001085; PRELIMINARY; PRT; 20 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE Bromelain (Fragment).
 OC Ananas comosus (Pineapple).
 OC Buxaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromelaceae; Ananas.
 OX NCBI_TAXID=4615;

RP SEQUENCE FROM N.A.
 RC STRAIN=CV_SMOOTH_CAYENNE; TISSUE=STEM;
 RA Robertson C.E., Goderough P.W.;
 RA "Cloning and expression of ananain gene from pineapple.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; A0009430; CAA08861.1; --.
 DR HSSP; P1408; IVAL.
 DR MEROPS; C01_026; --.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHProt_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_Protease_His; UNKNOWN_1.
 DR PROSITE; PS00639; THIOL_Protease_Asn; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 1 357 AA; 39606 MW; 0B393D035347C31 CRC64;
 SQ SEQUENCE 357 AA; 39606 MW; 0B393D035347C31 CRC64;
 Query Match 85.2%; Score 92; DB 10; Length 357;
 Best Local Similarity 94.4%; Pred. No. 7.8e-07;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OQ 1 VPOSIDWRYGAVTSVKNO 18
 Db 123 VPQSIDWRYGAVTSVKNO 140

RESULT 9
 O23800

SEQUENCE FROM N.A.
 STRAIN=CV_N67-10; TISSUE=FRUIT;
 Mutu E., Okamoto Y., Ota S.;
 "Cloning and sequencing of cysteine proteinases in Ananas comosus.";
 Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
 EMBL: D38533; BAA22545.1; -.
 HSSP: P07211; 1C1L.
 MEROPS: C01_028; -.
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; SHprt_acsite.
 DR Pfam: PF00112; Peptidase_C1; -.
 DR PRINTS: P000105; PAPAIN.
 DR PRODOM: PD000158; Peptidase_C1; -.
 DR PROSITE: PS000640; THIOL_PROTEASE ASN; 1.
 DR PROSITE: PS000139; THIOL_PROTEASE CYS; 1.
 DR PROSITE: PS000639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Hydrolase; Signal; Thiol protease.
 FT SIGNAL; 1 24 POTENTIAL.
 FT CHAIN 122 >340 FB22.
 FT NON_TER 340 340
 SQ SEQUENCE 340 AA; 37646 MW; 0055868006A903F CRC64;

Query Match 83.3%; Score 90; DB 10; Length 340;
 Best Local Similarity 85.0%; Pred. No. 1.6e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQ 20
 Db 123 VGGSIDWRYGAVTEVKDQN 142

RESULT 10
 O24641 PRELIMINARY; PRT; 352 AA.
 AC Q24641_05
 DR 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE FB31 precursor (FB3 Precursor) (Bromelain).
 OS Ananas comosus (Pineapple).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Bromeliaceae; Magnoliophyta; Liliopsida; Commelinidae incertae sedis;
 OC Bromeliaceae; Ananas.
 RN [1] _TaxID=4615;
 RP SEQUENCE FROM N.A.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR000705; PAPAIN.
 DR PRODOM: PD000158; Peptidase_C1; 1.
 DR PROSITE: PS000139; THIOL_PROTEASE ASN; 1.
 DR PROSITE: PS000639; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS000139; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Hydrolase; Signal; Thiol protease.
 FT SIGNAL; 1 24 POTENTIAL.
 FT CHAIN 123 >340 FB31 (FB13)
 SQ SEQUENCE 352 AA; 38823 MW; 5569AAC09BB3D367 CRC64;

Query Match 83.3%; Score 90; DB 10; Length 352;
 Best Local Similarity 85.0%; Pred. No. 1.6e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQ 20
 Db 124 VGGSIDWRYGAVTEVKDQN 143

RESULT 11
 Q95VA7 PRELIMINARY; PRT; 326 AA.
 ID Q95VA7
 AC Q95VA7
 DR 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE Cathepsin L.
 GN CAT-LIG.
 OS Fasciola gigantica.
 OC Echinostomida; Metacercaria; Platyhelminthes; Trematoda; Digenea;
 OC Fasciolidae; Echinostomatida; Fasciolidae; Fasciolidae; Fasciola.
 OC NCBI_TaxID=46835;
 RN [1]

SEQUENCE FROM N.A.
 RA Sobhon P., Meemon K., Grams R., Korge G., Hofmann A.;
 RT "Molecular cloning of cathepsin L encoding genes from Fasciola
 gigantica.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF119329; AAL23917.1; -
 DR InterPro: IPR000658; Peptidase_C1.
 DR InterPro: IPR000169; SHprt_acsite.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRODOM: PD000158; Peptidase_C1; 1.
 DR PROSITE: PS000640; THIOL_PROTEASE ASN; UNKNOWN_1.
 DR PROSITE: PS000139; THIOL_PROTEASE_CYS; UNKNOWN_1.
 DR PROSITE: PS000639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Hydrolase; Signal; Thiol protease.
 FT SIGNAL; 1 24 POTENTIAL.
 FT CHAIN 122 >340 FB22.
 SQ SEQUENCE 340 AA; 37646 MW; 0055868006A903F CRC64;

Query Match 83.3%; Score 90; DB 10; Length 340;
 Best Local Similarity 85.0%; Pred. No. 1.6e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQ 20
 Db 123 VGGSIDWRYGAVTEVKDQN 142

RESULT 12
 Q40922 PRELIMINARY; PRT; 454 AA.
 ID Q40922
 AC Q40922; Q40919;
 DR 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE PSEUDOTZAIN (Cysteine protease).
 GN PM33TSP
 OS Pseudotzain manziesii (Douglas-fir).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pseudotsuga.
 OC NCBI_TaxID=3357;
 RN [1] _TaxID=3357;
 RP SEQUENCE FROM N.A.

RC STRAIN=[MTRB] FRANCO;
 RX MEDLINE=96026908; PubMed=8682307;
 RA Tranbarger T.J.; Misra S.;
 RT structure and expression of a developmentally regulated cDNA encoding
 a cysteine protease (pseudozincin) from Douglas fir.;
 RT Gene 172:221-226 (1995).
 [2]
 RP SEQUENCE FROM N.A.
 RN STRAIN=[MTRB] FRANCO;
 RA Tranbarger T.J.; Misra S.;
 RL Physiol. Plantarum 0:0-0(0).
 RN [3]
 RP SEQUENCE OF 297-347 FROM N.A.
 RC TISSUE=WHOLE SPEDLING;
 RA Tranbarger T.J.; Misra S.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
 DR U41902; AAC44551; -.
 EMBL; Z49165; CAB8835.1; -.
 DR HSSP; P07711; ICUL.
 MEROPS; C01.029;
 InterPro; IPR000118; Granulin.
 InterPro; IPR000121; Peptidase_C1.
 InterPro; IPR000169; Short_acsive.
 InterPro; IPR00121; Peptidase_C1.
 InterPro; IPR00112; Peptidase_C1; 1.
 PRINTS; PR00705; PAPAIN.
 DR prodom; P0000158; Peptidase_C1; 1.
 SMART; SW00277; GRAN; 1.
 DR PROSITE; PS00640; THIOL_Protease ASN; 1.
 DR PROSITE; PS00139; THIOL_Protease CYS; 1.
 DR PROSITE; PS00639; THIOL_Protease HIS; 1.
 KW Hydrolase; Protease; Thiol_protease; Glycoprotein; lysosome; Zymogen; Signal;
 Sequence differentiation; developmental protein.
 DR MERO5; C01.022; -.
 DR InterPro; IPR000628; Peptidase_C1.
 DR InterPro; IPR000119; SHprt_acsive.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR prodom; P0000158; Peptidase_C1; 1.
 DR PROSITE; PS00540; THIOL_Protease ASN; 1.
 DR PROSITE; PS00039; THIOL_Protease HIS; 1.
 DR PROSITE; PS00639; THIOL_Protease HIS; 1.
 KW Hydrolase; Thiol_protease; Glycoprotein; lysosome; Zymogen; Signal;
 Sequence differentiation; developmental protein.
 DR SIGNAL; 17
 FT PROPEP 18 121 ACTIVATION PEPTIDE.
 FT CHAIN 122 294 CATHEPSIN L HEAVY CHAIN.
 FT PROPEP 295 298 BY SIMILARITY.
 FT CHAIN 299 339 CATHEPSIN L LIGHT CHAIN.
 FT ACT_SITE 145 146 BY SIMILARITY.
 FT ACT_SITE 285 285 BY SIMILARITY.
 FT ACT_SITE 306 306 BY SIMILARITY.
 FT DISUFRID 143 186 BY SIMILARITY.
 FT DISUFRID 177 219 BY SIMILARITY.
 FT DISUFRID 278 328 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 96 96 N LINKED (GlcNAc, . .) (POTENTIAL).
 SQ SEQUENCE 339 AA; 37847 MW; 7401F32812FDA33 CRC64;

Query Match 74.1%; Score 80; DB 10; Length 454;
 Best Local Similarity 73.7%; Pred. No. 9; e=0.05;
 Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VFOQSTDWRYGAVTVKQ 19
 DB 132 LPBSIDWREKGAVTVKQ 150

RESULT 13

Q26636 PRELIMINARY; PRT; 339 AA.
 ID Q26636 :
 AC O9STL5 ;
 DT 01-NOV-1996 (TREMBrel. 01, Created)
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)
 DB Cathepsin L precursor (EC 3.4.22.15).
 OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Sarcophagidae; Sarcophaga.
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
 MEDLINE=94253090; PubMed=8195162;
 RA Honma K.-I.; Kurata S.; Natori S.;
 RT "purification, characterization and cDNA cloning of procathepsin L
 from the culture medium of NIH-Sape-4, an embryonic cell line of
 Sarcophaga peregrina (flesh fly), and its involvement in the
 differentiation of imaginal discs";
 RT J. Biol. Chem. 269:15258-15264 (1994).
 RL FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN
 CC LYSOMES. REQUIRED FOR DIFFERENTIATION OF IMAGINAL DISKS.
 CC -!- CATALYTIC ACTIVITY: SPECIFICITY CLOSE TO THAT OF PAPAIN. AS
 COMPARED TO CATHESPIN B, CATHESPIN L EXHIBITS HIGHER ACTIVITY
 TOWARDS PROTEIN SUBSTRATES, BUT HAS LITTLE ACTIVITY ON Z-ARG-ARG-
 CC NMEC, AND NO PEPTIDYL-DIPEPTIDASE ACTIVITY.

RESULT 14

Q26636 PRELIMINARY; PRT; 364 AA.
 ID Q9STL5
 AC O9STL5 ;
 DT 01-MAY-2000 (TREMBrel. 13, Created)
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)
 DE Cysteine endopeptidase precursor-like protein.
 GN T29H11_130.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophytina; Magnoliophyta; eudicots; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Choisne N.; Robert C.; Brodtier P.; Wincker P.; Gattolico L.,
 RA Artiguenave R.; Saurin W.; Weissenbach J.; Mewes H.W.; Rudd S.,
 RA Lemcke K.; Mayer K.F.X.; Quertier F.; Salamoubat M.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
 DR HSSP; P07711; ICUL.
 MEROPS; C01.010;
 DR InterPro; IPR000886; ER_target.

DR InterPro; IPR000668; Peptidase_C1.
 DR IPR000169; SHprt_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00014; ER TARGET; UNKNOWN_1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; 1.
 DR Hydrolase; Thiol_protease.
 KW SEQUENCE: 364 AA; 40971 MW: 81C725D07DCB2E78 CRC64;
 SQ

Query Match 73.1%; Score 79; DB 10; Length 364;

Best Local Similarity 73.7%; Pred. No. 0.0001; Length 364;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDYGAVTSVKNQ 1.9
 Db 126 VPSSVDREKGAVTEVKNQ 1.44

RESULT 15

Q9GRW6 PRELIMINARY; PRT; 311 AA.
 ID Q9GRW6
 AC 09GRW6
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Procathepsin L3 (Fragment).
 OS Fasciola hepatica (Liver Fluke).
 OX Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 OC Echinostomida; Echinostomata; Fasciolidae; Fasciola.
 OC NCBITaxonID=6192;
 OX NCBI_TAXID=6192;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Harmsen M.M.; Cornelissen J.B.W.J.; van Milligen F.J.; Buijs H.E.C.M.,
 RA Jeurissen S.H., Boerema W.;
 RT "The propeptide of a cathepsin L-like proteinase induces protection
 against Fasciola hepatica infection in rats."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR AU279091; CAC12805..1.; -.
 HSPP; P07858; 1RUC.
 MEROPS: C01.033; -.

DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprt_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PRODOM; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE HIS; UNKNOWN_1.
 FT CHAIN 1 1 PROCATHEPSIN L3.
 SQ SEQUENCE 311 AA; 35409 MW; BB38A5DCB0510845 CRC64;

Query Match 71.3%; Score 77; DB 5; Length 311;

Best Local Similarity 73.7%; Pred. No. 0.0002; Length 311;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDYGAVTSVKNQ 1.9
 Db 93 VPASIDWRGYVTEVKDQ 1.11

Copyright (c) 1993 - 2002 Compugen Ltd.

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:07:14 ; Search time 13 Seconds

Title: US-09-674-738-2

perfect score: 105

Sequence: 1 VPOSIDWRDGSAGTYSVKNQG 20

Scoring table: BLOSUM62

Gapped 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% ; Listing first 45 summaries

Database : PIR_733,*

1: piel1:*

2: piel2:*

3: piel3:*

4: piel4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	105	100.0	20	S46204
2	105	100.0	345	T0839
3	94	89.5	20	S46205
4	93	88.6	212	T0364
5	93	88.6	356	T1014
6	93	88.6	356	T0740
7	88	83.8	454	JC4948
8	84	80.0	324	T1018
9	84	80.0	351	T10303
10	84	80.0	352	T0960
11	83	79.0	334	2
12	83	79.0	357	T0751
13	83	79.0	367	T12382
14	83	79.0	498	T3181
15	82	78.1	43	S35580
16	82	78.1	214	JN0718
17	82	78.1	334	1
18	82	78.1	339	A53810
19	82	78.1	357	2
20	82	78.1	360	T08122
21	81	77.1	361	2
22	81	77.1	364	T06707
23	81	77.1	368	JN0718
24	80	76.2	221	2
25	80	76.2	331	2
26	80	76.2	331	2
27	80	76.2	338	JC5442
28	80	76.2	356	2
29	80	76.2	361	T06708

DESCRIPTIONS

Query	Match	Local Similarity	Score	DB	Length	Pred.	No.	Mismatches	Indels	Gaps	Alignments
RESULT 1											
1 VPOSIDWRDGSAGTYSVKNQG 20											
C;Species: Ananas comosus (pineapple)											
C;Date: 16-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 07-May-1999											
R;Napper, A.D.; Bennett, S.P.; Borowski, M.; Holdridge, M.B.; Leonard, M.J.C.; Rogers, B											
Biochem. J. 301, 727-735, 1994											
A;Title: Purification and characterization of multiple forms of the pineapple-stem-deriv											
A;Reference number: S46204; MUID:94330946; PMID:8053898											
A;Accession: S46204											
A;Molecule type: protein											
A;Residues: 1-20 <NAP>											
C;Superfamily: papain											
C;Keywords: cysteine proteinase; hydrolase.											
Query Match Best Local Similarity 100.0%; Score 105; DB 2; Length 20;											
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
RESULT 2											
1 VPOSIDWRDGSAGTYSVKNQG 20											
C;Species: Ananas comosus (pineapple)											
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000											
R;Robertson, C.E.; Goodenough, P.W.											
submitted to the EMBL Data Library; November 1997											
A;Description: Cloning and expression of ananain gene from pineapple.											
A;Reference number: Z21616											
A;Accession: T07839											
A;Status: preliminary; translated from GB/EMBL/DDBJ											
A;Molecule type: mRNA											
A;Residues: 1-345 <R0>											
A;Cross-references: EMBL:AU02477; PIDN:CAA05487.1											
C;Genetics:											
A;Gene: AN1											
C;Superfamily: papain											
C;Keywords: cysteine proteinase; hydrolase											
F;1-24/Domain: signal sequence #status Predicted <MAT>											
F;25-345/Product: ananain #status Predicted <MAT>											
F;147-279/300/Active site: Cys, His, Asn #status Predicted											
Query Match 100.0%; Score 105; DB 2; Length 345;											
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											

cysteine proteins
oryzain (EC 3.4.22
cysteine Proteinas
cathepsin L (EC 3.
papain (EC 3.4.22.
cysteine proteins
cysteine proteins
drought-inducible
drought-inducible
probable cysteine
cysteine L-like c
cysteine proteins
carrot seed cysteai
cysteine proteins
cysteine proteins
cysteine proteins

Qy 1 VPOSIDWRDGSAGTYSVKNQ 20
Db 123 VPOSIDWRDGSAGTYSVKNQ 142

RESULT 3
S46205 comosain (EC 3.4.22.-) - Pineapple (fragment)
C:Species: Ananas comosus (pineapple)
C:Date: 16-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 07-May-1999
R:Napper, A.D.; Bennett, S.P.; Borowski, M.B.; Leonard, M.J.C.; Rogers, E.
Biochem. J. 301: 727-735, 1994
A;Title: Purification and characterization of multiple forms of the pineapple-stem-derivative
A;Reference number: S46204; MUID:94330946; PMID:8053898
A;Molecule type: protein
A;Residues: 1-20 <NAP>
C;Keywords: cysteine proteinase; hydrolase
C;Superfamily: papain
C;Residues: 1-20 <NAP>

Query Match Score 94; DB 2; Length 20;
Best Local Similarity 90.0%; Pred. No. 4.2e-09;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSAGTYSVKNQ 20
Db 1 VPOSIDWRDGSAGTYSVKNQ 20

RESULT 4
S03964 stem bromelain (EC 3.4.22.32) - pineapple
C:Species: Ananas comosus (pineapple)
C:Accession: S03964
;Riconja, A.; Rowan, A.D.; Buttle, D.J.; Rawlings, N.D.; Turk, V.; Barrett, A.J.
PEBS Lett. 247: 419-424; 1989
A;Title: Stem bromelain: amino acid sequence and implications for weak binding of crystal
A;Reference number: S03964; MUID:89232167; PMID:271443
A;Molecule type: protein
A;Residues: 1-212 <RIR>
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

Query Match Score 93; DB 2; Length 212;
Best Local Similarity 94.7%; Pred. No. 8.8e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSAGTYSVKNQ 19
Db 2 VPOSIDWRDGSAGTYSVKNQ 20

RESULT 5
T10514 Probable stem bromelain (EC 3.4.22.32) precursor - pineapple
N;Alternate names: FBSS precursor
C:Species: Ananas comosus (pineapple)
C:Accession: T10514
;R;Mut, E.; Okamoto, Y.; Ora, S.
submitted to the BMBL Data Library, October 1994
A;Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.
A;Accession: T10514
A;Molecule type: mRNA
A;Residues: 1-356 <MTUT>
A;Cross-references: EMBL:D38532
A;Experimental source: cv. N67-10; nearly mature fruit
C;Function:

A;Description: hydrolyzes proteins with broad specificity
C;Superfamily: papain
F;1-24/Domain: signal peptide; hydrolase
F;25-122/Domain: amino-terminal propeptide #status predicted <SIG>
F;123-339/Domain: stem bromelain #status predicted <MAT>
F;339-356/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;148-280/301/Active site: Cys, His, Asn #status predicted

Query Match Score 93; DB 2; Length 356;
Best Local Similarity 94.7%; Pred. No. 1.6e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSAGTYSVKNQ 19
Db 124 VPOSIDWRDGSAGTYSVKNQ 142

RESULT 6
T07840 ananain (EC 3.4.22.31) ANB precursor - pineapple
C:Species: Ananas comosus (pineapple)
C:Accession: T07840
;R;Robertson, C.E.; Goodenough, P.W.
submitted to the BMBL Data Library, July 1998
A;Description: Cloning and expression of ananain gene from pineapple.
A;Accession: T07840
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-356 <ROB>
A;Cross-references: EMBL:AJ009829; PIDN:CAA08860_1
C;Genetics:
A;Gen: ab
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

Query Match Score 93; DB 2; Length 356;
Best Local Similarity 94.7%; Pred. No. 1.6e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSAGTYSVKNQ 19
Db 124 VPOSIDWRDGSAGTYSVKNQ 142

RESULT 7
JC4848 cysteine proteinase (EC 3.4.22.-) - Douglas fir
N;Alternate names: pseudotruin
C:Species: Pseudotsuga menziesii (Douglas fir)
C:Accession: JC4848
;R;Traubacker, T.J.; Misra, S.
Gene 172, 221-226, 1996
A;Title: Structure and expression of a developmentally regulated cDNA encoding a cystein
A;Reference number: JC4848; MUID:96269408; PMID:8682307
A;Accession: JC4848
A;Molecule type: mRNA
A;Residues: 1-454 <TRA>
A;Cross-references: GB:U41902; NID:91208548; PID: AAC49455_1; PID:91208549
A;Note: it is uncertain whether Met-1, Met-15 or Met-41 is the initiator
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

Query Match Score 88; DB 2; Length 454;
Best Local Similarity 75.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPQSIDWRDGSAGTVSKVKNQ 20
 :|||||:|||||:|||||:
 Db 132 LPESIDWREKGAVTAKVNQ 151

RESULT 8

T10518 fruit bromelain (EC 3.4.22.33) FB1035 precursor - pineapple (fragment)

C;Species: Ananas comosus (pineapple)

C;Accession: T10518 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

R;Mutata, E.; Okamoto, Y.; Ota, S.
 A;Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.

A;Reference number: 217065

A;Accession: T10518

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-324 <MUT>

A;Cross-references: EMBL:D38534

A;Experimental source: cv. cv. N67-10; nearly mature fruit

C;Superfamily: papain

C;Keywords: cysteine proteinase; hydrolase

F;95-313/Product: fruit bromelain #status predicted <MAT>
 F;314-324/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F;120,352,273/Active site: Cys, His, Asn #status predicted

Query Match

Best Local Similarity 80.0%; Score 84; DB 2; Length 324;
 Matches 16; Conservative 84.2%; Pred. No. 4.4e-06; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPQSIDWRDGSAGTVSKVKNQ 19
 :|||||:|||||:|||||:
 Db 96 VPQSIDWRDGYGAVNEVKNQ 114

RESULT 9

T10503 fruit bromelain (EC 3.4.22.33) FB18 precursor - pineapple

C;Species: Ananas comosus (pineapple)

C;Accession: T10503 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

R;Mutata, E.; Aramaki, H.; Takata, Y.; Kono, A.; Okamoto, Y.; Ota, S.
 A;Description: Cloning and sequencing of fruit bromelain.

A;Reference number: 217060

A;Accession: T10503

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-351 <MUT>

A;Cross-references: EMBL:D14059

A;Experimental source: cv. N67-10; nearly mature fruit

C;Function: hydrolyses proteins; cysteine proteinase

A;Description: hydrolyses proteins; cysteine proteinase

C;Superfamily: papain

C;Keywords: cysteine proteinase; hydrolase

F;1-24/Domain: signal sequence #status predicted <SIG>
 F;1-24/Domain: amino-terminal propeptide #status predicted <PRO>
 F;122-339/Product: fruit bromelain #status predicted <MAT>
 F;340-351/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F;147,219,300/Active site: Cys, His, Asn #status predicted

Query Match

Best Local Similarity 80.0%; Score 84; DB 2; Length 352;
 Matches 16; Conservative 84.2%; Mismatches 3; Indels 0; Gaps 0;

QY 2 PQQSIDWRDGSAGTVSKVKNQ 20
 :|||||:|||||:
 Db 136 PQQSIDWRDAGVTPVTKNQ 154

RESULT 11

A58195 cathepsin L (EC 3.4.22.15) precursor - Pig

C;Species: Sus scrofa domesticus (domestic pig)

C;Accession: A58195; S59911; S59916

R;Okamura, N.; Tamba, M.; Uchiyama, Y.; Sugita, Y.; Dacheux, F.; Syntin, P.; Dacheux, J.
 Biochim Biophys Acta 1245, 221-226, 1995
 A;Title: Direct evidence for the elevated synthesis and secretion of procathepsin L in t
 A;Reference number: S59911; MUID:9606841; PMID:7495581

A;Accession: A58195

A;Molecule type: mRNA

A;Residues: 1-334 <OKR>

A;Cross-references: DDBJ:D37917; NID:9710655; PIDN:BAA07140.1; PID:gi1468964

A;Experimental source: epidemic

A;Note: the authors translated the codon CCG for residue 203 as Thr and CCC for residue 204 submitted to the DDBJ/EMBL/GenBank databases by Naomichi Okamura, 8 August 1994

A;Note: submitted to the DDBJ/EMBL/GenBank databases by Naomichi Okamura, 8 August 1994

A;Accession: S59911

A;Molecule type: protein

A;Residues: 18-31 <OK2>

C;Complex: heterodimer of disulfide linked chains produced from a single chain precursor

C;Function: catalyzes hydrolysis of peptide bonds in proteins

A;Description: protein degradation

A;Note: important role in the lysosomal degradation of proteins

C;Superfamily: papain

C;Keywords: cysteine proteinase; glycoprotein; heterodimer; hydrolase; lysosome; protein

F;1-17/Domain: signal sequence #status predicted <SIG>

RESULT 10

T09760

F:18-113/Domain: propeptide #status predicted <PRO>
 F:114-334/Domain: cathepsin L #status predicted <MAT>
 F:135-178:169-212,270-323/Disulfide bonds: #status predicted
 F:138,277,301/Active site: Cys, His, Asn #status predicted
 F:222,292/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match Score 83; DB 2; Length 334;
 Best Local Similarity 70.0%; Pred. No. 6.7e-06;
 Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGGAVTSVKNG 20
 Db 114 VPKSVDWREKGVTVAVKNG 133

RESULT 14
 Query Match Score 83; DB 2; Length 334;
 Best Local Similarity 70.0%; Pred. No. 6.7e-06;
 Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGGAVTSVKNG 20
 Db 126 VPRSIDWRVKGAVTPVKNG 145

hypothetical protein F41E6.6 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Accession: T31871
 R;Sammons, L.; Murray, J.
 submitted to the EMBL Data Library, July 1997
 A;Description: The sequence of C. elegans cosmid F41E6.
 A;Reference number: 221095
 A;Accession: T31871
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: DNA
 A;Residues: 1-498 <AM>
 A;Cross-references: EMBL:AF016448; PIDN:AA865956.1; GSPDB:GN000023; CESP:F41E6.6
 A;Experimental source: strain Bristol N2; clone F41E6
 C;Genetics:
 A;Gene: CESP; F41E6.6
 A;Map Position: 5
 A;Introns: 34/3; 102/3; 140/2; 180/2; 336/2; 357/2; 417/1

Query Match Score 83; DB 2; Length 498;
 Best Local Similarity 70.0%; Pred. No. 1e-05;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGGAVTSVKNG 20
 Db 264 LPESFDWREKGAVTPVKNG 283

RESULT 15
 Query Match Score 83; DB 2; Length 357;
 Best Local Similarity 88.9%; Pred. No. 7.2e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGGAVTSVK 18
 Db 123 VPOSIDWRNYGAVTSVK 140

cysteine proteinase (EC 3.4.22.-) IV - mountain papaya (fragment)
 C;Species: Carica pubescens (mountain papaya)
 C;Accession: S35580
 R;Malraeuens, V.; Jaziri, M.; van Beumen, J.; Schmek, A.G.; Kleinschmidt, T.; Lo
 Biol. Chem. Hoppe-Seyler 374, 501-506, 1993
 A;Title: Isolation and Preliminary Characterization of the cysteine-proteinases f
 A;Reference number: S355577; MUID:94030669; PMID:8216902
 A;Accession: S35580
 A;Molecule type: protein
 A;Residues: 1-43 <WAL>
 A;Note: the source is designated as *Carica carandaicensis*
 C;Superfamily: Papain
 C;Keywords: cysteine proteinase; hydrolase

Query Match Score 82; DB 2; Length 43;
 Best Local Similarity 78.9%; Pred. No. 9.8e-07;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDWRDGGAVTSVKNG 20
 Db 2 PEISDWRKGAVTPVKNG 20

Search completed: November 21, 2002, 16:10:31
 Job time : 13 secs

RESULT 13
 Query Match Score 83; DB 2; Length 195;
 Best Local Similarity 88.9%; Pred. No. 7.2e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGGAVTSVK 18
 Db 123 VPOSIDWRNYGAVTSVK 140

cysteine proteinase (EC 3.4.22.-) - common ice plant
 C;Species: Mesembryanthemum crystallinum (common ice plant)
 C;Accession: T12382
 R;Forsthoefel, N.R.; Cushman, J.C.
 A;Description: A novel sulphydryl endopeptidase is induced by environmental stress and p
 A;Reference number: 217502
 A;Accession: T12382
 A;Molecule type: mRNA
 A;Residues: 1-367 <PRO>
 A;Cross-references: EMBL:U30322; NID:9944915; PID:9944916
 A;Experimental source: leaves NaCl stressed; clone pSep7
 C;Genetics:
 A;Gene: Sep1
 C;Keywords: cysteine proteinase; hydrolase

Copyright (c) 1993 - 2002 Compugen Ltd.

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:07:14 ; Search time 13 Seconds

Title: US-09-674-738-1

Perfect score: 108

Sequence: 1 VPQSIDWRDYGAVTISKVQN 20

Scoring table: BLOSUM62

gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Listing first 45 summaries

No.	Score	Query Length	DB ID	Description
1	108	100.0	212 S03964	stem bromelain (EC 3.4.22.32) - pineapple
2	108	100.0	356 T0514	C;Species: Ananas comosus (pineapple)
3	108	100.0	356 T07840	C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change
4	99	91.7	324 T0517	R;Ritorto, A.; Rowan, A.D.; Buttle, D.J.; Rawlings, N.D.; Turk, V.; Barrett, A.J.
5	99	91.7	351 T10503	FEBs Lett. 247, 419-424, 1989
6	89.8	20	346205 S03964	A;Title: Stem bromelain: amino acid sequence and implications for weak binding of cystatin
7	93	86.1	20 S46204	A;Reference number: S03964; PMID:89232167; PMID:2714443
8	93	86.1	345 T07839	A;Accession: S03964
9	92	85.2	357 T07851	A;Molecule type: protein
10	90	83.3	340 T0516	A;Residues: I-212 <RT>
11	90	83.3	352 T0501	C;Superfamily: papain
12	80	45.4	2 JC4848	C;Keywords: cysteine proteinase; hydrolease
13	79	73.1	339 A33810	Query Match 100.0%; Score 108; DB 2; Length 212;
14	79	73.1	364 T06707	Best Local Similarity 100.0%; Pred. No. 7.3e-10;
15	78	72.2	368 JN0718	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
16	76	70.4	352 T09760	RESULT 2
17	76	70.4	363 S02149	probable stem bromelain (EC 3.4.22.32) precursor - pineapple
18	75	69.4	326 T09259	N;Alternate names: FBSB precursor
19	75	69.4	334 A38195	C;Species: Ananas comosus (pineapple)
20	75	69.4	358 JC7787	C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change
21	75	69.4	365 S02150	C;Accession: T10514
22	75	69.4	367 T12382	R;Muta, E.; Okamoto, Y.; Ota, S.
23	75	69.4	498 T021871	submitted to the EMBL Data library, October 1994
24	74	68.5	43 S035680	A;Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.
25	74	68.5	214 S46476	A;Reference number: 217065
26	74	68.5	334 1 KHS1	A;Accession: T10514
27	74	68.5	357 S49166	A;Status: preliminary; translated from GB/EMBL/DDJB
28	74	68.5	360 T08122	A;Molecule type: mRNA
29	74	68.5	361 2 BB4601	A;Residues: 1-356 <RT>
				A;Cross references: EMBL,D38532
				A;Experimental source: cv. N67-10; nearly mature fruit
				C;Function: Clotting factor Xa inhibitor
				A;Description: hydrolyzes proteins with broad specificity
				C;Superfamily: Papain
				C;Keywords: Cysteine proteinase; hydrolase
				F;1-24/Domain: signal sequence #status predicted <SIG>
				F;25-122/Domain: amino-terminal proptide #status predicted <PRO>
				F;123-338/Product: stem bromelain #status predicted <MTI>
				F;339-356/Domain: carboxyl-terminal proptide #status predicted <CTP>
				F;149,280/301/Active site: Cys, His, Asn #status predicted

SUMMARIES

ALIGNMENTS

Query Match 100.0%; Score 108; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

Qy 1 VPOSIDWRYGAVTSVKNQN 20
 Db 124 VPQSIDWRYGAVTSVKNQN 143

RESULT 3
 T07840
 ananain (EC 3.4.22.31) AN8 precursor - Pineapple
 C;Species: Ananas comosus (Pineapple)
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
 R;Robertson, C.E.; Goodenough, P.W.
 submitted to the EMBL Data Library, July 1998
 A;Description: Cloning and expression of ananain gene from pineapple.
 A;Reference number: Z16167
 A;Accession: T07840
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: mRNA
 A;Residues: 1-356 <ROB>
 A;Cross-references: EMBL:AJ009829; PIDN:CAA08860.1
 A;Gene: ananain
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 108; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

Qy 1 VPOSIDWRYGAVTSVKNQN 20
 Db 124 VPQSIDWRYGAVTSVKNQN 143

RESULT 4
 T10518
 fruit bromelain (EC 3.4.22.33) FB1035 precursor - pineapple (fragment)
 C;Species: Ananas comosus (Pineapple)
 C;Date: 16-Jul-1998 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 R;Mutu, E.; Okamoto, Y.; Ota, S.
 submitted to the EMBL Data Library, October 1994
 A;Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.
 A;Accession: T10518
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: mRNA
 A;Residues: 1-324 <MUT>
 A;Cross-references: EMBL:D38534
 A;Experimental source: cv. N67-10; nearly mature fruit
 C;Keywords: cysteine proteinase; hydrolase

Query Match 91.7%; Score 99; DB 2; Length 324;
 Best Local Similarity 90.0%; Pred. No. 3.3e-08; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2;

Qy 1 VPOSIDWRYGAVTSVKNQN 20
 Db 96 VPQSIDWRYGAVTSVKNQN 115

RESULT 5
 T10503
 fruit bromelain (EC 3.4.22.33) FB18 precursor - pineapple

C;Species: Ananas comosus (Pineapple)
 C;Accession: T10503
 R;Mutu, E.; Aramaki, H.; Takara, Y.; Kono, A.; Okamoto, Y.; Ota, S.
 submitted to the EMBL Data Library, January 1993
 A;Description: Cloning and sequencing of fruit bromelain.
 A;Accession: T10503
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: mRNA
 A;Residues: 1-351 <MUT>
 A;Cross-references: ENBL:D14059
 A;Experimental source: cv. N67-10; nearly mature fruit
 C;Function:
 A;Description: hydrolyses proteins; cysteine proteinase
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase

Query Match 91.7%; Score 99; DB 2; Length 351;
 Best Local Similarity 90.0%; Pred. No. 3.6e-08; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQN 20
 Db 123 VPQSIDWRYGAVTSVKNQN 142

RESULT 6
 S46205
 comosain (EC 3.4.22.-) - pineapple (fragment)
 C;Species: Ananas comosus (Pineapple)
 C;Date: 16-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 07-May-1999
 R;Napper, A.D.; Bennett, S.P.; Borowski, M.; Holdridge, M.B.; Leonard, M.J.C.; Rogers, E.
 Biochem. J. 301, 727-735, 1994
 A;Title: Purification and characterization of multiple forms of the pineapple-stem-deriv.
 A;Accession: S46205
 A;Molecule type: protein
 A;Residues: 1-120 <NAP>
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase

Query Match 89.8%; Score 97; DB 2; Length 20;
 Best Local Similarity 94.7%; Pred. No. 3.1e-09; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQN 19
 Db 1 VPQSIDWRYGAVTSVKNQN 19

RESULT 7
 S46204
 ananain (EC 3.4.22.31) - pineapple (fragment)
 C;Species: Ananas comosus (Pineapple)
 C;Date: 16-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 07-May-1999
 C;Accession: S46204
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase

Query Match 91.7%; Score 99; DB 2; Length 324;
 Best Local Similarity 90.0%; Pred. No. 3.3e-08; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKNQN 20
 Db 96 VPQSIDWRYGAVTSVKNQN 115

Query Match 86.1%; Score 93; DB 2; Length 20;
 Best Local Similarity 94.7%; Pred. No. 1.4e-08; Indels 0; Mismatches 1; Gaps 0;

Matches 18; Conservative 0; Indels 0; Gaps 0;

RESULT 8
 T07839
 Query 1 VPQSIDWRDYGAVTISKNO 19
 Database 1 VPQSIDWRDSGAVTISKNO 19

anain (EC 3.4.22.31) precursor - pineapple
 C;Species: Ananas comosus (pineapple)
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
 C;Accession: T07839
 R;Robertson, C.E.; Goodenough, P.W.
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-345 <ROB>
 A;Cross-references: EMBL:AU02477; PIDN:CAA05487.1
 C;Genetics:
 A;Gene: AN1
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase
 F;1-24/Domain: signal sequence #status predicted <SIG>
 Query Match 86.1%; Score 93; DB 2; Length 345;
 Best Local Similarity 94.7%; Pred. No. 3.3e-07; Indels 0; Mismatches 1; Gaps 0;
 Matches 18; Conservative 0; Indels 0; Gaps 0;

Query 1 VPQSIDWRDYGAVTISKNO 19
 Database 123 VPQSIDWRDSGAVTISKNO 141

RESULT 9
 T07851
 anain (EC 3.4.22.31) precursor AN11 - pineapple
 N;Alternate names: cysteine proteinase precursor AN11
 C;Species: Ananas comosus (pineapple)
 C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
 C;Accession: T07851
 C;Robertson, C.E.; Goodebaugh, P.W.
 R;Robertson, C.E.; Goodebaugh, P.W.
 submitted to the EMBL Data Library, July 1998
 A;Description: Cloning and expression of anain gene from pineapple.
 A;Reference number: Z16167
 A;Accession: T07851
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-357 <ROB>
 A;Cross-references: EMBL:AJ009830; PIDN:CAA08861.1
 C;Genetics:
 A;Gene: an1
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase
 F;1-26/Domain: signal sequence #status predicted <SIG>
 F;27-357/Domain: product: anain AN11 #status predicted <MAT>
 F;147-281/302/Active site: Cys, His, Asn #status predicted
 Query Match 85.2%; Score 92; DB 2; Length 357;
 Best Local Similarity 94.4%; Pred. No. 5e-07; Indels 0; Mismatches 1; Gaps 0;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 83.3%; Score 90; DB 2; Length 352;
 Best Local Similarity 85.0%; Pred. No. 1e-06; Indels 0; Mismatches 2; Gaps 0;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Query 1 VPQSIDWRDYGAVTISKNO 20
 Database 124 VPQSIDWRDYGAVTISKNO 143

RESULT 10
 T10516
 fruit bromelain (EC 3.4.22.33) FB22 precursor - pineapple (fragment)
 C;Species: Ananas comosus (pineapple)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10516
 R;Mutta, E.; Okamoto, Y.; Ota, S.
 submitted to the EMBL Data Library, October 1994
 A;Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.
 A;Reference number: Z17065
 A;Accession: T10516
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-340 <MTU>
 A;Cross-references: EMBL:D38533
 A;Experimental source: CV. N67-10; nearly mature fruit
 C;Function:
 A;Description: hydrolyzes proteins with broad specificity; cysteine proteinase
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-121/Domain: amino-terminal propeptide #status predicted <PRO>
 F;122-339/Product: fruit bromelain #status predicted <PRO>
 F;147-280/301/Active site: Cys, His, Asn #status predicted
 Query Match 83.3%; Score 90; DB 2; Length 340;
 Best Local Similarity 85.0%; Pred. No. 9.9e-07; Indels 0; Mismatches 1; Gaps 0;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query 1 VPQSIDWRDYGAVTEVKDN 20
 Database 123 VQOSIDWRDYGAVTEVKDN 142

RESULT 11
 T10501
 fruit bromelain (EC 3.4.22.33) FB13 precursor - pineapple
 C;Species: Ananas comosus (pineapple)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10501
 R;Mutta, E.; Aramaki, H.; Takata, Y.; Kono, A.; Okamoto, Y.; Ota, S.
 submitted to the EMBL Data Library, July 1993
 A;Description: Cloning and sequencing of fruit bromelain.
 A;Reference number: Z17060
 A;Accession: T10501
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-352 <MTU>
 A;Cross-references: EMBL:D14058
 A;Experimental source: CV. N67-10; fruit; nearly mature
 C;Function:
 A;Description: hydrolyzes proteins; cysteine proteinase
 C;Superfamily: papain
 C;Keywords: cysteine proteinase; hydrolase
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-122/Domain: amino-terminal propeptide #status predicted <PRO>
 F;133-340/Product: fruit bromelain #status predicted <PRO>
 F;341-352/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F;148-281/302/Active site: Cys, His, Asn #status predicted
 Query Match 83.3%; Score 90; DB 2; Length 352;
 Best Local Similarity 85.0%; Pred. No. 1e-06; Indels 0; Mismatches 2; Gaps 0;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Query 1 VPQSIDWRDYGAVTISKNO 20
 Database 124 VPQSIDWRDYGAVTISKNO 143

RESULT 12
 JC4448
 cysteine proteinase (EC 3.4.22.-) - Douglas fir

N;Alternate names: pseudozain
C;Species: *Pseudotsuga menziesii* (Douglas fir)
C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 04-Feb-2000
C;Accession: JC4848
R;Tranbarger, T.J.; Misra, S.
A;Title: Structure and expression of a developmentally regulated cDNA encoding a cystein
Gene 172, 221-226, 1996
A;Cross-references: GB:U41902; NID:91208548; PIDN:AC49455.1; PMID:8682107
A;Accession: JC4848
A;Molecule type: mRNA
A;Residues: 1-454 <TRA>
A;Note: it is uncertain whether Met-1, Met-15 or Met-41 is the initiator
C;Comment: This enzyme catalyzes the initial stages of storage protein mobilization during
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

A;Introns: 148/1; 274/3
C;Superfamily: papain
C;Species: *cysteine proteinase; hydrolase*
F;150;287;308/Active site: Cys, His, Asn #status predicted

Query Match 74.1%; Score 80; DB 2; Length 454;
Best Local Similarity 73.7%; Pred. No. 5.6e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VPOSIDWRYGAVTSVKHQ 19
Db 132 LPESIDWREKGAVTAVKQ 150

RESULT 13

A53810 L (EC 3.4.22.15) precursor - flesh fly (*Sarcophaga peregrina*)
C;Species: *Sarcophaga peregrina*
C;Accession: A53810
R;Honma, K.; Kurata, S.; Natori, S.
J. Biol. Chem. 269, 15258-15264, 1994
A;Title: Purification, characterization, and cDNA cloning of procathepsin L from the cuticle differentiation of imaginal discs
A;Reference number: A53810; PMID:94253090; PMID:8195162
A;Accession: A53810
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-339 <HOM>
A;Cross-references: GB:D16533; NID:9505139; PIDN:BAA03970.1; PMID:9505140
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

F;146;285;306/Active site: Cys, His, Asn #status predicted
Query Match 73.1%; Score 79; DB 2; Length 339;
Best Local Similarity 68.4%; Pred. No. 5.9e-05;
Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VPOSIDWRYGAVTSVKHQ 19
Db 122 VPKSVDWREHGAVTGVRDQ 140

RESULT 14

T06707 cysteine proteinase (EC 3.4.22.-) T29H11.130 - Arabidopsis thaliana
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Accession: T06707
R;Querier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigau
submitted to the Protein Sequence Database, April 1999
A;Reference number: 215793
A;Accession: T06707
A;Molecule type: DNA
A;Residues: 1-36 <QUE>
A;Cross-references: EMBL:AL049659; GSPDB:GN00061; ATSP:T29H11.130
A;Experimental source: cultivar Columbia; BAC clone T29H11
C;Genetics:
A;Gene: ATSP:T29H11.130
A;Map position: 3

Query Match 72.2%; Score 78; DB 2; Length 368;
Best Local Similarity 68.4%; Pred. No. 9.3e-05;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VPOSIDWRYGAVTSVKHQ 19
Db 135 LPEDFDORDHQAEGAVTEVKHQ 153

Search completed: November 21, 2002, 16:10:31
Job time : 14 secs

GenCore version 5.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:07:39 ; Search time 11.5 Seconds
 (without alignments)
 51.170 Million cell updates/sec

Title: US-09-674-738-1
 Perfect score: 108
 Sequence: I VPOSIDWRDYGAIVTSVKNON 20

Scoring table: BLOSIM62
 GapOp 10.0 , GapExt 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
 Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata1/1/aa/5A COMB.pep: *
 2: /cgn2_6/ptodata1/1/aa/5B COMB.pep: *
 3: /cgn2_6/ptodata1/1/aa/6A COMB.pep: *
 4: /cgn2_6/prodata1/1/aa/6B COMB.pep: *
 5: /cgn2_6/prodata1/1/aa/PCUTUS COMB.pep: *
 6: /cgn2_6/ptodata1/1/aa/backfiles.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

*** SUMMARIES**

Result No.	Score	Query	Match length	DB ID	Description
1	108	100.0	213	2	US-09-578-701A-1
2	108	100.0	213	3	US-09-360-693-1
3	88	81.0	18	4	US-09-382-689A-1
4	79	73.1	127	4	US-09-325-932A-176
5	79	73.1	148	4	US-09-325-932A-167
6	79	73.1	278	4	US-09-325-932A-165
7	79	73.1	339	3	US-09-120-365-62
8	79	73.1	339	3	US-09-120-365-80
9	79	73.1	339	4	US-09-515-039-62
10	79	73.1	339	4	US-09-515-039-80
11	75	69.4	443	3	US-08-883-526-4
12	75	69.4	443	4	US-09-120-665-69
13	75	69.4	443	4	US-09-151-039-69
14	74	68.5	334	3	US-09-120-665-64
15	74	68.5	334	4	US-09-515-039-64
16	73	67.6	396	4	US-09-325-932A-153
17	72	66.7	129	4	US-09-325-932A-159
18	72	66.7	191	3	US-08-906-69-178
19	72	66.7	191	4	US-08-639-675A-178
20	72	66.7	191	4	US-09-004-731-77
21	72	66.7	191	4	US-09-012-331-177
22	66.7	191	4	US-08-749-699-77	
23	72	66.7	191	4	US-08-012-692-178
24	72	66.7	191	4	US-08-606-63-178
25	72	66.7	191	4	US-09-004-729-95
26	72	66.7	220	4	US-09-004-731-89
27	66.7	220	4	US-09-004-731-95	

RESULT 1
 US-08-578-701A-1
 Application US/08578701A
 Patent No. 5824305

GENERAL INFORMATION:

APPLICANT: MYOTT, Tracy L.
 TITLE OF INVENTION: MEDICAL USE OF STEM BROMELAIN PROTEASE

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penne & Edwards
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/578,701A
 FILING DATE: 22-DEC-1995
 CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Baldwin, Geraldine F.
 REGISTRATION NUMBER: 31.232
 REFERENCE/DOCKET NUMBER: 7116-046

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-7741/8864
 TELEX: 66111 PENNIE

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 213 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-578-701A-1

Query Match 100.0%; Score 108; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred. No. 9.5e-11; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; MisMatch 0;

Qy 1 VPOSIDWRDYGAIVTSVKNON 20
 Db 2 VPOSIDWRDYGAIVTSVKNON 21

RESULT 2
US-08-360-693-1
; Sequence 1, Application US/08360693
; Patent No. 5926640
; GENERAL INFORMATION:
; APPLICANT: MYNOTT, Tracey L.
; TITLE OF INVENTION: USE OF ENZYMES, ESPECIALLY BROMELAIN, IN
; NUMBER OF SEQUENCES: 1
; TITLE OF INVENTION: THE TREATMENT OF DIARRHOEA
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/360.693
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/01374
; FILING DATE: 30-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 931389.4
; FILING DATE: 25-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9308164.4
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9213862.7
; FILING DATE: 30-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30357/115/KIST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5339
; TELEX: 90136
; INQUIRY FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-360-693-1
Query Match Score 108; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 9.5e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 1 VPQSIDWRDYGAVTSVKQN 20
Db 2 VPQSIDWRDYGAVTSVKQN 21

RESULT 3
US-09-382-689A-1
; Sequence 1, Application US/09382689A
; Patent No. 6335427
; GENERAL INFORMATION:
; APPLICANT: Mynott, Tracy Leanne
; APPLICANT: Engwarda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623-07600.01
; CURRENT APPLICATION NUMBER: US/09/382,689A

RESULT 2
US-09-360-693-1
; CURRENT FILING DATE: 1999-08-25
; PRIORITY APPLICATION NUMBER: PCT/GB98/00592
; PRIOR FILING DATE: 1998-02-25
; PRIORITY APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIORITY APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIORITY APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-382-689A-1
Query Match Score 88; DB 4; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
Qy 1 VPQSIDWRDYGAVTSVKN 18
Db 1 VPQSIDWRDYGAVNEVKN 18
;
RESULT 4
US-09-325-932A-176
; Sequence 176, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; death and their use in the modification of forestry plant develo
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-325-932A-176
Query Match Score 79; DB 4; Length 127;
Best Local Similarity 68.4%; Pred. No. 4.3e-06;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
;
Qy 1 VPQSIDWRDYGAVTSVKNO 19
Db 67 LPKDFDMRDHGAVTAKNQ 85
;
RESULT 5
US-09-325-932A-167
; Sequence 167, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; death and their use in the modification of forestry plant develo
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
;

US-09-325-932A-167

Query Match

Score 79; DB 4; Length 148;

Best Local Similarity 68.4%; Pred. No. 5.2e-06;

Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Flinn, Barry

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant develo

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325, 932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO: 80

LENGTH: 339

TYPE: PRT

ORGANISM: *Bucaluptus grandis*

RESULT 6

US-09-325-932A-165

Query Match

Score 79; DB 4; Length 148;

Best Local Similarity 68.4%; Pred. No. 5.2e-06;

Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Lasham, Annette

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant develo

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325, 932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO: 165

LENGTH: 278

TYPE: PRT

ORGANISM: *Bucaluptus grandis*

RESULT 7

US-09-120-365-62

Query Match

Score 79; DB 4; Length 278;

Best Local Similarity 68.4%; Pred. No. 1.2e-05;

Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Flinn, Barry

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant develo

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325, 932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO: 165

LENGTH: 278

TYPE: PRT

ORGANISM: *Bucaluptus grandis*

RESULT 8

US-09-120-365-80

Query Match

Score 79; DB 3; Length 339;

Best Local Similarity 68.4%; Pred. No. 1.5e-05;

Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Natori, Shunji

TITLE OF INVENTION: NEW PROTEASE

FILE REFERENCE: 32290-144749

CURRENT APPLICATION NUMBER: US/09/515, 039

CURRENT FILING DATE: 2000-03-06

EARLIER APPLICATION NUMBER: JP 9-333 474

EARLIER FILING DATE: 1997-11-18

NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 62

LENGTH: 339

TYPE: PRT

ORGANISM: *Sarcophaga cathepsin L*

RESULT 9

US-09-515-039-62

Query Match

Score 79; DB 3; Length 339;

Best Local Similarity 68.4%; Pred. No. 1.5e-05;

Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Natori, Shunji

TITLE OF INVENTION: NEW PROTEASE

FILE REFERENCE: 32290-144749

CURRENT APPLICATION NUMBER: US/09/515, 039

CURRENT FILING DATE: 2000-03-06

EARLIER APPLICATION NUMBER: JP 9-333 474

EARLIER FILING DATE: 1997-11-18

NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 62

LENGTH: 339

TYPE: PRT

ORGANISM: *Sarcophaga cathepsin L*

RESULT 10

US-09-515-039-80

Query Match

Score 79; DB 3; Length 339;

Best Local Similarity 68.4%; Pred. No. 1.5e-05;

Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Natori, Shunji

TITLE OF INVENTION: NEW PROTEASE

FILE REFERENCE: 32290-144749

CURRENT APPLICATION NUMBER: US/09/515, 039

CURRENT FILING DATE: 2000-03-05

EARLIER APPLICATION NUMBER: JP 9-333 474

EARLIER FILING DATE: 1997-11-18

NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 60

LENGTH: 339

TYPE: PRT

ORGANISM: *Sarcophaga cathepsin L*

RESULT 11

US-09-515-039-80

Query Match

Score 79; DB 3; Length 339;

Best Local Similarity 68.4%; Pred. No. 1.5e-05;

Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Natori, Shunji

TITLE OF INVENTION: NEW PROTEASE

FILE REFERENCE: 32290-144749

CURRENT APPLICATION NUMBER: US/09/515, 039

CURRENT FILING DATE: 2000-03-05

EARLIER APPLICATION NUMBER: JP 9-333 474

EARLIER FILING DATE: 1997-11-18

NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 60

LENGTH: 339

TYPE: PRT

ORGANISM: *Sarcophaga cathepsin L*

Query Match 73.1%; Score 79; DB 4; Length 339;
 Best Local Similarity 68.4%; Pred. No. 1.5e-05;
 Matches 13; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKNQ 19
 Db 122 VPKSVDWRERHGAVTGVKDQ 140

RESULT 11
 Sequence 4, Application US/08883526

Patent No. 6033883
 GENERAL INFORMATION:
 APPLICANT: Bandhan, Olga
 APPLICANT: Guegler, Kari J.
 APPLICANT: Corley, Neil C.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: NEW HUMAN CATEPSIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastaSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/883,526
 FILING DATE: Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: P-0331 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 334 amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: 1468964
 CLONE: GenBank
 US-08 883 526 -4

Query Match 69.4%; Score 75; DB 3; Length 334;
 Best Local Similarity 68.4%; Pred. No. 7e-05;
 Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKNQ 19
 Db 114 VPKSVDWRERHGAVTGVKNQ 132

RESULT 12
 US-09-120-365-64
 Sequence 69, Application US/09120365
 Patent No. 610354
 GENERAL INFORMATION:

APPLICANT: Natori, Shunji
 TITLE OF INVENTION: NEW PROTEASE
 FILE REFERENCE: 32290-144749
 CURRENT APPLICATION NUMBER: US/09/120,365
 CURRENT FILING DATE: 1998-07-22
 EARLIER APPLICATION NUMBER: JP 9-333 474
 EARLIER FILING DATE: 1997-11-18
 NUMBER OF SEQ ID NOS: 101
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 69
 LENGTH: 443
 TYPE: PRT
 ORGANISM: Leishmania
 US-09-120-365-64

Query Match 69.4%; Score 75; DB 3; Length 443;
 Best Local Similarity 68.4%; Pred. No. 9.4e-05;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKNQ 19
 Db 126 VPDAVDWREKGAVTPVKNQ 144

RESULT 13
 US-09-515-039-69
 Sequence 69, Application US/09515039
 Patent No. 6214399
 GENERAL INFORMATION:
 APPLICANT: Natori, Shunji
 TITLE OF INVENTION: NEW PROTEASE
 FILE REFERENCE: 32290-144749
 CURRENT APPLICATION NUMBER: US/09/515,039
 CURRENT FILING DATE: 2000-03-06
 EARLIER APPLICATION NUMBER: JP 9-333 474
 EARLIER FILING DATE: 1997-11-18
 NUMBER OF SEQ ID NOS: 101
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 69
 LENGTH: 443
 TYPE: PRT
 ORGANISM: Leishmania
 US-09-515-039-69

Query Match 69.4%; Score 75; DB 4; Length 443;
 Best Local Similarity 68.4%; Pred. No. 9.3e-05;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKNQ 19
 Db 126 VPDAVDWREKGAVTPVKNQ 144

RESULT 14
 US-09-120-365-64
 Sequence 64, Application US/09120365
 Patent No. 6103514
 GENERAL INFORMATION:
 APPLICANT: Natori, Shunji
 TITLE OF INVENTION: NEW PROTEASE
 FILE REFERENCE: 32290-144749
 CURRENT APPLICATION NUMBER: US/09/120,365
 CURRENT FILING DATE: 1998-07-22
 EARLIER APPLICATION NUMBER: JP 9-333 474
 EARLIER FILING DATE: 1997-11-18
 NUMBER OF SEQ ID NOS: 101
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 64
 LENGTH: 334
 TYPE: PRT
 ORGANISM: Murine cathepsin L
 US-09-120-365-64

Query Match 68.5%; Score 74; DB 3; Length 334;
 Best Local Similarity 63.2%; Pred. No. 0.0001; Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDYGAVTSVKHQ 19
 Db 114 IPKSVDWRKGCVTPVKHQ 132

RESULT 15

US-09-515-039-64

; Sequence 64, Application US/09515039

; Patent No. 6214599

; GENERAL INFORMATION:

; APPLICANT: Natori, Shunji

; TITLE OF INVENTION: NEW PROTEASE

; FILE REFERENCE: 32200-14749

; CURRENT APPLICATION NUMBER: US/09/515, 039

; EARLIER APPLICATION NUMBER: JP 9-333 474

; EARLIER FILING DATE: 1997-11-18

; NUMBER OF SEQ ID NOs: 101

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 64

; LENGTH: 334

; TYPE: PRT

; ORGANISM: Murine cathepsin L

; US-09-515-039-64

Query Match 68.5%; Score 74; DB 4; Length 334;

Best Local Similarity 63.2%; Pred. No. 0.0001; Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDYGAVTSVKHQ 19
 Db 114 IPKSVDWRKGCVTPVKHQ 132

Search completed: November 21, 2002, 16:11:01
 Job time : 12.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:07:39 ; Search time 11.5 Seconds

(without alignments) 51.170 Million cell updates/sec

Title: US-09-674-738-2
Perfect score: 105
Sequence: 1 VPQSIDWRSAGAVTSVKNG 20

Scoring table: BIOSUM62

Gapop 10.0 , gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0 %

Maximum Match 100 %

Listing first 45 summaries

Database : Issued_Patents_AA,*

1: /cgn2_6/prodata1/1;aa/5A COMB.pep: *
2: /cgn2_6/prodata1/1;aa/5B COMB.pep: *
3: /cgn2_6/prodata1/1;aa/6A COMB.pep: *
4: /cgn2_6/prodata1/1;aa/6B COMB.pep: *
5: /cgn2_6/prodata1/1;aa/PCTUS COMB.pep: *
6: /cgn2_6/prodata1/1;aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query Score Match Length DB ID Description

Result No.	Score	Match Length	DB	ID	Description
1	93	89.5	213	2	US-09-578-701A-1
2	93	88.6	213	2	US-09-360-693-1
3	83	79.0	334	3	US-09-883-526-4
4	83	79.0	443	3	US-09-120-665-69
5	83	79.0	443	4	US-09-515-039-69
6	82	78.1	127	4	US-09-325-932A-176
7	82	78.1	148	4	US-09-325-932A-167
8	82	78.1	278	4	US-09-325-932A-165
9	82	78.1	334	3	US-09-120-365-64
10	82	78.1	334	4	US-09-515-039-64
11	82	78.1	339	3	US-09-120-665-62
12	82	78.1	339	4	US-09-515-039-62
13	82	78.1	339	4	US-09-515-039-80
14	81	77.1	396	4	US-09-325-932A-153
15	81	77.1	129	4	US-09-325-932A-159
16	80	76.2	442	4	US-09-515-039-66
17	80	76.2	443	4	US-09-821-994-65
18	79	75.2	18	4	US-09-821-994-65
19	79	75.2	212	4	US-08-860-255A-4
20	79	75.2	225	4	US-08-821-994-51
21	79	75.2	345	3	US-09-120-365-73
22	79	75.2	345	4	US-09-515-039-73
23	79	75.2	345	4	US-08-821-994-67
24	79	75.2	370	4	US-09-325-932A-169
25	79	75.2	374	4	US-08-821-994-67
26	79	75.2	374	4	US-08-821-994-67
27	79	75.2	374	4	US-08-821-994-67

ALIGNMENTS

RESULT 1
US-09-578-701A-1
Sequence 1, Application US/08578701A
Patient No. 5824305
GENERAL INFORMATION:
APPLICANT: MYNOT, TRACY L.
TITLE OF INVENTION: MEDICAL USE OF STEM BROMELAIN PROTEASE
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penrie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/578-701A
FILING DATE: 22-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31-232
REFERENCE/DOCKET NUMBER: 7116-046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-578-701A-1

Query Match 89.6%; Score 93; DB 2; Length 213;
Best Local Similarity 94.7%; Pred. No. 1.1e-07; 1; Indels 0; Gaps 0;
Matches 18; Conservative 1; Sequence 1, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 10, Appli
Sequence 11, Appli
Sequence 12, Appli
Sequence 13, Appli
Sequence 14, Appli
Sequence 15, Appli
Sequence 16, Appli
Sequence 17, Appli
Sequence 18, Appli
Sequence 19, Appli
Sequence 20, Appli
Sequence 21, Appli
Sequence 22, Appli
Sequence 23, Appli
Sequence 24, Appli
Sequence 25, Appli
Sequence 26, Appli
Sequence 27, Appli
Sequence 28, Appli
Sequence 29, Appli
Sequence 30, Appli
Sequence 31, Appli
Sequence 32, Appli
Sequence 33, Appli
Sequence 34, Appli
Sequence 35, Appli
Sequence 36, Appli
Sequence 37, Appli
Sequence 38, Appli
Sequence 39, Appli
Sequence 40, Appli
Sequence 41, Appli
Sequence 42, Appli
Sequence 43, Appli
Sequence 44, Appli
Sequence 45, Appli
Sequence 46, Appli
Sequence 47, Appli
Sequence 48, Appli
Sequence 49, Appli
Sequence 50, Appli
Sequence 51, Appli
Sequence 52, Appli
Sequence 53, Appli
Sequence 54, Appli
Sequence 55, Appli
Sequence 56, Appli
Sequence 57, Appli
Sequence 58, Appli
Sequence 59, Appli
Sequence 60, Appli
Sequence 61, Appli
Sequence 62, Appli
Sequence 63, Appli
Sequence 64, Appli
Sequence 65, Appli
Sequence 66, Appli
Sequence 67, Appli
Sequence 68, Appli
Sequence 69, Appli
Sequence 70, Appli
Sequence 71, Appli
Sequence 72, Appli
Sequence 73, Appli
Sequence 74, Appli
Sequence 75, Appli
Sequence 76, Appli
Sequence 77, Appli
Sequence 78, Appli
Sequence 79, Appli
Sequence 80, Appli
Sequence 81, Appli
Sequence 82, Appli
Sequence 83, Appli
Sequence 84, Appli
Sequence 85, Appli
Sequence 86, Appli
Sequence 87, Appli
Sequence 88, Appli
Sequence 89, Appli
Sequence 90, Appli
Sequence 91, Appli
Sequence 92, Appli
Sequence 93, Appli
Sequence 94, Appli
Sequence 95, Appli
Sequence 96, Appli
Sequence 97, Appli
Sequence 98, Appli
Sequence 99, Appli
Sequence 100, Appli

RESULT 2
US-08-360-693-1
; Sequence 1, Application US/08360693
; Patent No. 5928640
; GENERAL INFORMATION:
; APPLICANT: MYNORT, Tracey L.
; TITLE OF INVENTION: USE OF ENZYMES, ESPECIALLY BROMELAIN, IN
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,693
; FILING DATE: 14-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/01374
; FILING DATE: 30-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 931389.4
; FILING DATE: 25-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9308164.4
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9213862.7
; FILING DATE: 30-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; SPANNEDNESS: single
; TOPOLOGY: linear
; US-08-360-693-1

Query Match 88.6%; Score 93; DB 2; Length 213;
Best Local Similarity 94.7%; Pred. No. 1.1e-07; Mismatches 0; Gaps 0;
Matches 18; Conservative 0; Indels 1;

Qy 1 VPOSIDWRDSGAVTSVKNG 19
Db 2 VPQSIDWRDGYGAVTSVKNG 20

RESULT 3
US-08-883-526-4
; Sequence 4, Application US/08883526
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW HUMAN CATHEPSIN
; NUMBER OF SEQUENCES: 4

RESULT 2
US-08-360-693-1
; Sequence 1, Application US/08360693
; Patent No. 5928640
; GENERAL INFORMATION:
; APPLICANT: Incyte Pharmaceuticals, Inc.
; ADDRESS: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,526
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0331 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; SPANNEDNESS: single
; TOPOLOGY: Linear
; IMMEDIATE SOURCE:
; LIBRARY: 1468964
; CLONE: GenBank
; US-08-883-526-4

Query Match 79.0%; Score 83; DB 3; Length 334;
Best Local Similarity 70.0%; Prd. No. 7.8e-06; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDSGAVTSVKNG 20
Db 114 VPQSIDWRDGYGAVTSVKNG 133

RESULT 4
US-09-120-365-69
; Sequence 69, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Leishmania
; US-09-120-365-69

Query Match 79.0%; Score 83; DB 3; Length 443;
Best Local Similarity 70.0%; Prd. No. 1.1e-05; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDSGAVTSVKNG 20

```

Db 126 VPDAVDREKGAVTPVKNQG 145
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 167
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-325-932A-167

Query Match 78.1%; Score 82; DB 4; Length 148;
Best Local Similarity 70.0%; Pred. No. 4.3e-06;
Matches 14; Conservative 70.0%; 3; Mismatches 3; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/325,932A
EARLIER APPLICATION NUMBER: JP 9-333 474
NUMBER OF SEQ ID NOS: 101
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 69
LENGTH: 443
TYPE: PRT
ORGANISM: Leishmania
US-09-515-039-69

Query Match 79.0%; Score 83; DB 4; Length 443;
Best Local Similarity 70.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 70.0%; 3; Mismatches 3; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/325,932A
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 176
LENGTH: 127
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-09-325-932A-176

Query Match 78.1%; Score 82; DB 4; Length 127;
Best Local Similarity 70.0%; Pred. No. 3.6e-06;
Matches 14; Conservative 70.0%; 3; Mismatches 3; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/325,932A
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 176
LENGTH: 127
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-09-325-932A-176

Query Match 78.1%; Score 82; DB 4; Length 278;
Best Local Similarity 70.0%; Pred. No. 9.1e-05;
Matches 14; Conservative 70.0%; 2; Mismatches 4; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/325,932A
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 165
LENGTH: 278
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-09-325-932A-165

Query Match 78.1%; Score 82; DB 4; Length 278;
Best Local Similarity 70.0%; Pred. No. 9.1e-05;
Matches 14; Conservative 70.0%; 2; Mismatches 4; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/325,932A
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 165
LENGTH: 278
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-09-325-932A-165

Query Match 78.1%; Score 82; DB 4; Length 334;
Best Local Similarity 65.0%; Pred. No. 1.1e-05;
Matches 13; Conservative 70.0%; 4; Mismatches 3; Indels 0; Gaps 0;
TITLE OF INVENTION: Compositions affecting programmed cell
death and their use in the modification of forestry plant develo
FILE REFERENCE: 1032
CURRENT APPLICATION NUMBER: US/09/325,932A
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
CURRENT FILING DATE: 1999-06-04

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 80
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
; US-09-120-365-80

Query Match 78.1%; Score 82; DB 3; Length 339;
Best Local Similarity 70.0%; Pred. No. 1.2e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Organism: Sarcophaga cathepsin L

Qy 1 VPOSIDWRSQAVTSVKNOQ 20
Db 122 VPKSVDREHGAVTGVKDQG 141

RESULT 13
US-09-515-039-62
; Sequence 62, Application US/09515039
; Patent No. 621159
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515, 039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 64
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Murine cathepsin L
; US-09-515-039-64

Query Match 78.1%; Score 82; DB 4; Length 334;
Best Local Similarity 65.0%; Pred. No. 1.1e-05;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Organism: Murine cathepsin L

Qy 1 VPOSIDWRSQAVTSVKNOQ 20
Db 114 IPKSVDWREHGCVTPVKNOQ 133

RESULT 11
US-09-120-365-62
; Sequence 62, Application US/09120365
; Patent No. 610514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120, 365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 62
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
; US-09-120-365-62

Query Match 78.1%; Score 82; DB 3; Length 339;
Best Local Similarity 70.0%; Pred. No. 1.2e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Organism: Sarcophaga cathepsin L

Qy 1 VPOSIDWRSQAVTSVKNOQ 20
Db 122 VPKSVDREHGAVTGVKDQG 141

RESULT 12
US-09-120-365-80
; Sequence 80, Application US/09120365
; Patent No. 610514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120, 365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 80
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
; US-09-515-039-80

Query Match 78.1%; Score 82; DB 4; Length 339;
Best Local Similarity 70.0%; Pred. No. 1.2e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Organism: Sarcophaga cathepsin L

Qy 1 VPOSIDWRSQAVTSVKNOQ 20
Db 122 VPKSVDREHGAVTGVKDQG 141

```

RESULT 15
US-09-325-932A-153
; Sequence 153, Application US/09325932A
; Patient No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develo
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325, 932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 153
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Pinus radiata
; US-09-325-932A-153

Query Match 77.1%; Score 81; DB 4; Length 396;
Best Local Similarity 70.0%; Pred. No. 2e-05;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VPGSIDWRDSDGAVTSVKNQG 20
DB 160 LPESIDWREKGAVNPVKNQG 179

Search completed: November 21, 2002, 16:11:01
Job time : 11.5 secs

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:08:44 ; Search time 137.5 seconds
(without alignments)

93.779 Million cell updates/sec

Title: US-09-674-738-1

Perfect score: 108

Sequence: 1 VPQSIDWRDYGAVTISVKQN 20

Scoring table: BILOSUM62

Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main: *

1: /cgn2_6/prodata/1/paa/US084 COMB.pep:*

2: /cgn2_6/prodata/1/paa/US05 COMB.pep:*

3: /cgn2_6/prodata/1/paa/US07 COMB.pep:*

4: /cgn2_6/prodata/1/paa/US080 COMB.pep:*

5: /cgn2_6/prodata/1/paa/US081 COMB.pep:*

6: /cgn2_6/prodata/1/paa/US082 COMB.pep:*

7: /cgn2_6/prodata/1/paa/US084 COMB.pep:*

8: /cgn2_6/prodata/1/paa/US085 COMB.pep:*

9: /cgn2_6/prodata/1/paa/US086 COMB.pep:*

10: /cgn2_6/prodata/1/paa/US087 COMB.pep:*

11: /cgn2_6/prodata/1/paa/US088 COMB.pep:*

12: /cgn2_6/prodata/1/paa/US089 COMB.pep:*

13: /cgn2_6/prodata/1/paa/US090 COMB.pep:*

14: /cgn2_6/prodata/1/paa/US091 COMB.pep:*

15: /cgn2_6/prodata/1/paa/US092 COMB.pep:*

16: /cgn2_6/prodata/1/paa/US093 COMB.pep:*

17: /cgn2_6/prodata/1/paa/US094 COMB.pep:*

18: /cgn2_6/prodata/1/paa/US095 COMB.pep:*

20: /cgn2_6/prodata/1/paa/US096 COMB.pep:*

21: /cgn2_6/prodata/1/paa/US097 COMB.pep:*

22: /cgn2_6/prodata/1/paa/US098 COMB.pep:*

23: /cgn2_6/prodata/1/paa/US099 COMB.pep:*

24: /cgn2_6/prodata/1/paa/US100 COMB.pep:*

25: /cgn2_6/prodata/1/paa/US101 COMB.pep:*

26: /cgn2_6/prodata/1/paa/US102 COMB.pep:*

27: /cgn2_6/prodata/1/paa/US102 COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
US-09-674-738-1

; Sequence 1, Application US/09674738

; GENERAL INFORMATION:

; APPLICANT: URSAPHARM Arzneimittel GmbH

; TITLE OF INVENTION: Use of Bromelain Proteases for Inhibiting Blood Coagulation

; FILE REFERENCE: 80054

; CURRENT APPLICATION NUMBER: US/09/674,738

; CURRENT FILING DATE: 2000-10-31

; PRIORITY APPLICATION NUMBER: PCT/EP98/04406

; PRIORITY FILING DATE: 1998-07-15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO: 1

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Pine-apple (Bromeliaceae)

; US-09-674-738-1

Result No.	Score	% Match Length	DB ID	Description
1	108	100.0	20	US-09-674-738-1
2	108	100.0	21	US-09-382-685A-2
3	108	100.0	21	US-09-382-688A-3
4	108	100.0	21	US-09-750-210A-3
5	108	100.0	212	US-08-244-796-15
6	100.0	356	21	US-09-791-537-84882

Query Match Score 108; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPQSIDWRDYGAVTISVKQN 20
Do 1 VPQSIDWRDYGAVTISVKQN 20

RESULT 2
US-09-382-685A-2
; Sequence 2, Application US/09382685A
; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne
; APPLICANT: Engewarda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623_0750001
; CURRENT APPLICATION NUMBER: US/09/382,685A
; PRIOR APPLICATION NUMBER: PCT/GB98/00591
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119-6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252-7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850-9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827-7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-382-685K-2

Query Match 100.0%; Score 108; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.3e-10; Mismatches 0;
Matches 20; Conservative 0; Gaps 0;

Qy 1 VPQSIDWRDYGAVTSVKNQN 20
Db 2 VPQSIDWRDYGAVTSVKNQN 21

RESULT 3
US-09-382-688-3
; Sequence 3, Application US/09382688A
; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne
; APPLICANT: Engewarda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623_0740001
; CURRENT APPLICATION NUMBER: US/09/382,688A
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119-6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252-7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850-9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827-7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-382-688-3

Query Match 100.0%; Score 108; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.3e-10; Mismatches 0;
Matches 20; Conservative 0; Gaps 0;

Qy 1 VPQSIDWRDYGAVTSVKNQN 20
Db 2 VPQSIDWRDYGAVTSVKNQN 21

RESULT 4
US-09-750-210A-3
; Sequence 3, Application US/09750210A
; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne
; APPLICANT: Engewarda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623_0740001
; CURRENT APPLICATION NUMBER: US/09/750-210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119-6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252-7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850-9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827-7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-3

Query Match 100.0%; Score 108; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.3e-10; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQSIDWRDYGAVTSVKNQN 20
Db 2 VPQSIDWRDYGAVTSVKNQN 21

RESULT 5
US-08-244-796-15
; Sequence 15, Application US/08244796
; GENERAL INFORMATION:
; APPLICANT: ANDREW R. MILNER, MICHAEL PANACCIOL, TERENCE W.
; TITLE OF INVENTION: LIVER FLUKE VACCINE AND POLYPEPTIDES USEFUL FOR SAME
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA, GARDEN CITY
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,796
; FILING DATE: 15-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL7109
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: PCT/AU94/00051
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLO, FRANK
; REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9316
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (1516) 742 4343
 TELEFAX: (1516) 742 4366
 INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
 LENGTH: 212 amino acids
 TYPE: amino acid
 STRANDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein

RESULT 6
 US-09-791-537-84879
 Sequence 84879, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; PELLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 15055
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 84879
 ; LENGTH: 356
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus

Query Match 100.0%; Score 108; DB 6; Length 212;
 Best Local Similarity 100.0%; Pred. No. 9.1e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQSIDWDRYGAIVTSVKQN 20
 Db 2 VPQSIDWDRYGAIVTSVKQN 21

RESULT 7
 US-09-791-537-84882
 Sequence 84882, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 84882
 ; LENGTH: 356
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus

Query Match 100.0%; Score 108; DB 6; Length 212;
 Best Local Similarity 100.0%; Pred. No. 9.1e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQSIDWDRYGAIVTSVKQN 20
 Db 2 VPQSIDWDRYGAIVTSVKQN 21

RESULT 8
 US-09-743-987-16
 Sequence 16, Application US/09743987
 ; GENERAL INFORMATION:
 ; APPLICANT: Myrott, Tracey Lehanne
 ; PELLICANT: Crossett, Ben
 ; APPLICANT: Myrott, Tracey Lehanne
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM PINEAPPLE STEM
 ; FILE REFERENCE: 0628.0980000
 ; CURRENT APPLICATION NUMBER: US/09/743, 987
 ; CURRENT FILING DATE: 2001-08-01
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02784
 ; PRIOR FILING DATE: 1999-08-24
 ; PRIOR APPLICATION NUMBER: GB 9819138
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 311
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus

Query Match 91.7%; Score 99; DB 21; Length 311;
 Best Local Similarity 90.0%; Pred. No. 4e-07; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPQSIDWDRYGAIVTSVKQN 20
 Db 96 VPQSIDWDRYGAIVNEVKQN 115

RESULT 9
 US-09-743-987-20
 Sequence 20, Application US/09743987
 ; GENERAL INFORMATION:
 ; APPLICANT: Myrott, Tracey Lehanne
 ; APPLICANT: Crossett, Ben
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM PINEAPPLE STEM
 ; FILE REFERENCE: 0628.0980000
 ; CURRENT APPLICATION NUMBER: US/09/743, 987
 ; CURRENT FILING DATE: 2001-08-01
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02784
 ; PRIOR FILING DATE: 1999-08-24
 ; PRIOR APPLICATION NUMBER: GB 9819138
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus

Query Match 91.7%; Score 99; DB 21; Length 324;
 Best Local Similarity 90.0%; Pred. No. 4.1e-07; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPQSIDWDRYGAIVTSVKQN 20
 Db 96 VPQSIDWDRYGAIVNEVKQN 115

RESULT 10
 US-09-791-537-69324
 Sequence 69324, Application US/09791537
 ; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69324
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-791-537-69324
Query Match 91.7%; Score 99; DB 21; Length 324;
Best Local Similarity 90.0%; Pred. No. 4.1e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 VPOSIDWRDYGAVTSVKNQ 20
Db 96 VPQSIDWRDYGAVNEVKNQ 115

RESULT 13
US-09-791-537-69324-4
Sequence 4, Application US/09382685A
; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne
; APPLICANT: Engwenda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623-0750001
; CURRENT APPLICATION NUMBER: US/09/382 685A
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00591
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus
US-09-382-685A-4

Query Match 89.8%; Score 97; DB 17; Length 20;
Best Local Similarity 94.7%; Pred. No. 4.2e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDYGAVTSVKNQ 19
Db 1 VPQSIDWRNNGAVTSVKNQ 19

RESULT 14
US-09-791-537-688-5
Sequence 5, Application US/09382688A
; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne
; APPLICANT: Engwenda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623-0740001
; CURRENT APPLICATION NUMBER: US/09/382 688A
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus
US-09-791-537-84889

Query Match 91.7%; Score 99; DB 21; Length 351;
Best Local Similarity 90.0%; Pred. No. 4.5e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 VPOSIDWRDYGAVTSVKNQ 20
Db 123 VPQSIDWRDYGAVNEVKNQ 142

RESULT 12
US-09-791-537-84889
Sequence 84889, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84889
; LENGTH: 351
; TYPE: PRT

Mon Nov 25 08:17:50 2002

```

; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Ananas comosus
; US-09-382-688-5

Query Match 89.8%; Score 97; DB 17; Length 20;
Best Local Similarity 94.7%; Pred. No. 4.2e-08; Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPQSIDWWRDGAVTTSVKNQ 19
Db 1 VPQSIDWWRNYGAVTTSVKNQ 19

RESULT 15
US-09-750-210A-5
Sequence 5, Application US/09750210A
GENERAL INFORMATION:
APPLICANT: Myrott, Tracy Lehanne
APPLICANT: Engwerda, Christian
APPLICANT: Peak, Keith
TITLE OF INVENTION: Component of Bromelain
FILE REFERENCE: 0523 074001
CURRENT APPLICATION NUMBER: US/09/750,210A
CURRENT FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: US 09/382,688
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: PCT/GB98/00590
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: GB 9706119.6
PRIOR FILING DATE: 1997-03-25
PRIOR APPLICATION NUMBER: GB 9704252.7
PRIOR FILING DATE: 1997-02-28
PRIOR APPLICATION NUMBER: GB 9703850.9
PRIOR FILING DATE: 1997-02-25
PRIOR APPLICATION NUMBER: GB 9703827.7
PRIOR FILING DATE: 1997-02-25
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 20
TYPE: PRT
ORGANISM: Ananas comosus
US-09-750-210A-5

Query Match 89.8%; Score 97; DB 21; Length 20;
Best Local Similarity 94.7%; Pred. No. 4.2e-08; Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPQSIDWWRDGAVTTSVKNQ 19
Db 1 VPQSIDWWRNYGAVTTSVKNQ 19

```

Search completed: November 21, 2002, 16:15:42
 Job time : 138.5 secs



GenCore version 5.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:08:44 ; Search time 137.5 Seconds
 (without alignments)
 93.779 Million cell updates/sec

Title: US-09-674-738-2

Perfect score: 105

Sequence: 1 VFOSIDWDRSGAVTSMKNGQ 20

Scoring table: BLOSUM62

Gapext 0.0 , Gapekt 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

database : Pending Patents AA Main:*

1: /cgnd_6/prodata/1/paa/PCTUS COMB.pep:*

2: /cgnd_6/prodata/1/paa/US06_7COMB.pep:*

3: /cgnd_6/prodata/1/paa/US07_7COMB.pep:*

4: /cgnd_6/prodata/1/paa/US08_7COMB.pep:*

5: /cgnd_6/prodata/1/paa/US08_7COMB.pep:*

6: /cgnd_6/prodata/1/paa/US08_7COMB.pep:*

7: /cgnd_6/prodata/1/paa/US083 COMB.pep:*

8: /cgnd_6/prodata/1/paa/US084 COMB.pep:*

9: /cgnd_6/prodata/1/paa/US085 COMB.pep:*

10: /cgnd_6/prodata/1/paa/US087 COMB.pep:*

11: /cgnd_6/prodata/1/paa/US088 COMB.pep:*

12: /cgnd_6/prodata/1/paa/US089 COMB.pep:*

13: /cgnd_6/prodata/1/paa/US090 COMB.pep:*

14: /cgnd_6/prodata/1/paa/US091 COMB.pep:*

15: /cgnd_6/prodata/1/paa/US092 COMB.pep:*

16: /cgnd_6/prodata/1/paa/US093 COMB.pep:*

17: /cgnd_6/prodata/1/paa/US094 COMB.pep:*

18: /cgnd_6/prodata/1/paa/US095 COMB.pep:*

19: /cgnd_6/prodata/1/paa/US096 COMB.pep:*

20: /cgnd_6/prodata/1/paa/US097 COMB.pep:*

21: /cgnd_6/prodata/1/paa/US098 COMB.pep:*

22: /cgnd_6/prodata/1/paa/US099 COMB.pep:*

23: /cgnd_6/prodata/1/paa/US100 COMB.pep:*

24: /cgnd_6/prodata/1/paa/US101 COMB.pep:*

25: /cgnd_6/prodata/1/paa/US102 COMB.pep:*

26: /cgnd_6/prodata/1/paa/US103 COMB.pep:*

27: /cgnd_6/prodata/1/paa/US60_7COMB.pep:*

7 94 89.5 20 17 US-09-382-685A-4

8 94 89.5 20 17 US-09-382-688-5

9 94 89.5 20 21 US-09-750-210A-5

10 93 88.6 20 20 US-09-674-738-1

11 93 88.6 21 17 US-09-382-685A-2

12 93 88.6 21 17 US-09-382-688-3

13 93 88.6 21 21 US-09-750-210A-3

14 93 88.6 21 19 US-09-743-987-16

15 93 88.6 20 21 US-09-791-537-84879

16 93 88.6 21 17 US-09-791-537-84882

17 88 83.8 454 21 US-09-791-537-32256

18 85 81.0 334 21 US-09-791-537-84891

19 84 80.0 311 21 US-09-743-987-20

20 84 80.0 324 21 US-09-791-537-69324

21 84 80.0 324 21 US-09-791-537-69324

22 84 80.0 351 21 US-09-743-987-14

23 84 80.0 351 21 US-09-791-537-84889

24 83 79.0 326 26 US-09-791-537-45992

25 83 79.0 334 19 US-09-791-537-84892

26 83 79.0 334 21 US-09-791-537-84895

27 83 79.0 334 21 US-09-791-537-84896

28 83 79.0 357 21 US-09-791-537-91358

29 83 79.0 374 21 US-09-791-537-91358

30 83 79.0 399 26 US-10-219-939-60863

31 83 79.0 399 27 US-60-324-109-29193

32 83 79.0 443 19 US-09-519-283-4

33 83 79.0 443 19 US-09-791-537-82405

34 83 79.0 443 21 US-09-791-537-61257

35 83 79.0 498 21 US-09-791-537-34918

36 83 79.0 498 27 US-60-610-039-6416

37 82 78.1 127 17 US-09-327-373-176

38 82 78.1 127 23 US-09-900-004A-176

39 82 78.1 127 26 US-10-219-220-176

40 82 78.1 148 17 US-09-337-313-167

41 82 78.1 148 23 US-09-004A-167

42 82 78.1 148 26 US-10-219-220-167

43 82 78.1 212 21 US-09-731-537-13192

44 82 78.1 212 21 US-09-337-313-165

45 82 78.1 278 23 US-09-990-004A-165

ALIGNMENTS

RESULT 1
 US-09-382-685A-3
 ; Sequence 3, Application US/09362685A
 ; GENERAL INFORMATION:
 ; APPLICANT: Munott, Tracy Leanne
 ; APPLICANT: Engwerda, Christian
 ; APPLICANT: Peak, Keith
 ; TITLE OF INVENTION: Component of Bromeain
 FILE REFERENCE: 0623_075001
 CURRENT APPLICATION NUMBER: US/09/382-685A
 CURRENT FILING DATE: 1999-08-25
 PRIOR APPLICATION NUMBER: PCT/GB98/00591
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: GB 9706119.6
 PRIOR APPLICATION NUMBER: GB 9704252.7
 PRIOR FILING DATE: 1997-01-28
 PRIOR APPLICATION NUMBER: GB 9703850.9
 PRIOR FILING DATE: 1997-03-25
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 20
 TYPE: PRT
 ORGANISM: Ananas comosus
 US-09-382-685A-3

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	105	100.0	20	17 US-09-382-685A-3
2	105	100.0	20	17 US-09-382-688-4
3	105	100.0	20	20 US-09-674-738-2
4	105	100.0	20	21 US-09-750-210A-4
5	105	100.0	21	US-09-791-537-848408
6	105	100.0	345	21 US-09-791-537-84883
7	100.0	21	US-09-791-537-84883	

Query Match 100.0%; Score 105; DB 17; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.7e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 4
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

RESULT 4
 US-09-750-210A-4
 ; Sequence 4, Application US/09750210A
 ; GENERAL INFORMATION:
 ; APPLICANT: Mynott, Tracy Leanne
 ; APPLICANT: Engwarda, Christian
 ; APPLICANT: Peek, Keith
 ; TITLE OF INVENTION: Component of Bromelain
 ; FILE REFERENCE: 0623.0740001
 ; CURRENT APPLICATION NUMBER: US/09/750,210A
 ; CURRENT FILING DATE: 2000-12-29
 ; PRIOR APPLICATION NUMBER: US 09/382,688
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: PCT/GB98/005590
 ; PRIOR FILING DATE: 1999-02-25
 ; PRIOR APPLICATION NUMBER: GB 9706119.6
 ; PRIOR FILING DATE: 1997-03-25
 ; PRIOR APPLICATION NUMBER: GB 9704252.7
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: GB 9706119.6
 ; PRIOR FILING DATE: 1997-03-25
 ; PRIOR APPLICATION NUMBER: GB 9704252.7
 ; PRIOR FILING DATE: 1997-02-28
 ; PRIOR APPLICATION NUMBER: GB 9703850.9
 ; PRIOR FILING DATE: 1997-02-25
 ; PRIOR APPLICATION NUMBER: GB 9703827.7
 ; PRIOR FILING DATE: 1997-02-25
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 4
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus
 US-09-362-688-4

Query Match 100.0%; Score 105; DB 17; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.7e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 5
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

RESULT 5
 US-09-791-537-82408
 ; Sequence 5, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; FILE REFERENCE: 261.210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 82408
 ; LENGTH: 216
 ; TYPE: PRT
 ; ORGANISM: Ananas comosus
 US-09-791-537-82408

Query Match 100.0%; Score 105; DB 21; Length 216;
 Best Local Similarity 100.0%; Pred. No. 4.9e-08; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 6
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

Query Match 100.0%; Score 105; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.7e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 3
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

Query Match 100.0%; Score 105; DB 21; Length 216;
 Best Local Similarity 100.0%; Pred. No. 4.9e-08; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 2
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

Query Match 100.0%; Score 105; DB 21; Length 216;
 Best Local Similarity 100.0%; Pred. No. 4.9e-08; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Result 1
 Db 1 VPQSTDWRDGSAGTTSVKNQG 20

US-09-791-537-84883

; Sequence 84883, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE

; FILE REFERENCE: 261210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO: 84883

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Ananas comosus

US-09-791-537-84883

; Query Match 100.0%; Score 105; DB 21; Length 345;

; Best Local Similarity 100.0%; Pred. No. 8.2e-08; Mismatches 0; Indels 0; Gaps 0;

; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Qy 1 VPOSIDWRDGSAGATSVKNG 20

; Db 123 VPOSIDWRDGSAGATSVKNG 142

RESULT 7 US-09-382-685A-4

; Sequence 4, Application US/09382685A

; GENERAL INFORMATION:

; APPLICANT: Myrott, Tracy Lehanne

; APPLICANT: Engwarda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; FILE REFERENCE: 0623.0750001

; CURRENT APPLICATION NUMBER: US/09/382,685A

; CURRENT FILING DATE: 1999-08-25

; PRIOR APPLICATION NUMBER: PCT/GB98/00591

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-03-25

; PRIOR APPLICATION NUMBER: GB 9704252.7

; PRIOR FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 4

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Ananas comosus

US-09-382-685A-4

; Query Match 100.0%; Score 105; DB 21; Length 345;

; Best Local Similarity 100.0%; Pred. No. 8.2e-08; Mismatches 0; Indels 0; Gaps 0;

; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Qy 1 VPOSIDWRDGSAGATSVKNG 20

; Db 123 VPOSIDWRDGSAGATSVKNG 20

RESULT 8 US-09-382-685B-5

; Sequence 84885, Application US/09382688A

; GENERAL INFORMATION:

; APPLICANT: Myrott, Tracy Lehanne

; APPLICANT: Engwarda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; FILE REFERENCE: 0623.0750210A

; CURRENT APPLICATION NUMBER: US/09750210A

; CURRENT FILING DATE: 1999-08-25

; PRIOR APPLICATION NUMBER: PCT/GB98/00590

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: GB 9706119.6

; PRIOR FILING DATE: 1997-03-25

; PRIOR APPLICATION NUMBER: GB 9704252.7

; PRIOR FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

RESULT 9 US-09-750-210A-5

; Sequence 5, Application US/09750210A

; GENERAL INFORMATION:

; APPLICANT: Myrott, Tracy Lehanne

; APPLICANT: Engwarda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; FILE REFERENCE: 0623.0740001

; CURRENT APPLICATION NUMBER: US/09/750,210A

; CURRENT FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: US 09/382,688

; PRIOR FILING DATE: 1999-08-25

; PRIOR APPLICATION NUMBER: PCT/GB98/00590

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: GB 9706119.6

; PRIOR FILING DATE: 1997-03-25

; PRIOR APPLICATION NUMBER: GB 9704252.7

; PRIOR FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 5

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Ananas comosus

US-09-750-210A-5

; Query Match 89.5%; Score 94; DB 21; Length 20;

; Best Local Similarity 90.0%; Pred. No. 2.1e-07; Mismatches 1; Indels 0; Gaps 0;

; Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

; Qy 1 VPOSIDWRDGSAGATSVKNG 20

; Db 1 VPOSIDWRDGSAGATSVKNG 20

RESULT 10 US-09-674-738-1

; Sequence 10, Application US/09382688A

; GENERAL INFORMATION:

; APPLICANT: Engwarda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; FILE REFERENCE: 0623.0740001

; CURRENT APPLICATION NUMBER: GB 9704252.7

; CURRENT FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 1

; LENGTH: 20

RESULT 10 US-09-674-738-1

; Sequence 10, Application US/09382688A

; GENERAL INFORMATION:

; APPLICANT: Engwarda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; FILE REFERENCE: 0623.0740001

; CURRENT APPLICATION NUMBER: GB 9704252.7

; CURRENT FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 1

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Ananas comosus

US-09-674-738-1

; Query Match 89.5%; Score 94; DB 21; Length 20;

; Best Local Similarity 90.0%; Pred. No. 2.1e-07; Mismatches 1; Indels 0; Gaps 0;

; Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

; Qy 1 VPOSIDWRDGSAGATSVKNG 20

; Db 1 VPOSIDWRDGSAGATSVKNG 20

Sequence 1; Application US/09674738
; GENERAL INFORMATION:
; APPLICANT: URSAPHARM Arzneimittel GmbH
; TITLE OF INVENTION: Use of Bromelain Proteases for Inhibiting Blood Coagulation
; TITLE OF INVENTION: Coagulation
; FILE REFERENCE: 80054
; CURRENT APPLICATION NUMBER: US/09/674,738
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: PCT/EP98/044406
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: pine-apple (Bromeliaceae)
US-09-674-738-1

Query Match 88.6%; Score 93; DB 20; Length 20;
Best Local Similarity 94.7%; Pred. No. 3.1e-07; Mismatches 1; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

US-09-382-685A-3

RESULT 11
US-09-382-685A-2

; GENERAL INFORMATION:
; Sequence 2, Application US/09382685A

; APPLICANT: Mynot, Tracy Lehanne

; APPLICANT: Engwerda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; CURRENT APPLICATION NUMBER: US/09/382,685A

; CURRENT FILING DATE: 1998-08-25

; PRIOR APPLICATION NUMBER: PCT/GB98/00591

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: GB 9706119.6

; PRIOR FILING DATE: 1997-03-25

; PRIOR APPLICATION NUMBER: GB 9704252.7

; PRIOR FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 2
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus

US-09-382-685A-2

Query Match 88.6%; Score 93; DB 21; Length 21;
Best Local Similarity 94.7%; Pred. No. 3.2e-07; Mismatches 1; Indels 0; Gaps 0;

US-09-382-685-3

RESULT 12
US-09-382-688-3

; Sequence 3, Application US/09382688A

; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne

; APPLICANT: Engwerda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; CURRENT APPLICATION NUMBER: US/09/382,688A

; CURRENT FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus

RESULT 13
US-09-750-210A-3

; Sequence 3, Application US/09750210A

; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne

; APPLICANT: Engwerda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

; CURRENT APPLICATION NUMBER: US/09/750,210A

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/382,688

; PRIOR FILING DATE: 1999-08-25

; PRIOR APPLICATION NUMBER: PCT/GB98/00590

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: GB 9706119.6

; PRIOR FILING DATE: 1997-01-25

; PRIOR APPLICATION NUMBER: GB 9704252.7

; PRIOR FILING DATE: 1997-02-28

; PRIOR APPLICATION NUMBER: GB 9703850.9

; PRIOR FILING DATE: 1997-02-25

; PRIOR APPLICATION NUMBER: GB 9703827.7

; PRIOR FILING DATE: 1997-02-25

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus

RESULT 14
US-08-244-736-15

; Sequence 15, Application US/08244736

; GENERAL INFORMATION:
; APPLICANT: Mynot, Tracy Lehanne

; APPLICANT: Engwerda, Christian

; APPLICANT: Peak, Keith

; TITLE OF INVENTION: Component of Bromelain

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2002, 16:09:09 ; Search time 8 Seconds
(without alignments)

86.116 Million cell updates/sec

Title: US-09-674-738-1

Perfect score: 108

Sequence: 1 VPOSTDWRYGAVTSVKQ 20

Scoring table: BLASTm62

Gapext 10.0 , Gapext 0.5

Searched: 99223 seqs, 34446456 residues

Total number of hits satisfying chosen parameters: 9923

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%; Maximum Match 100%
Listing first 45 summaries

Database : Pending Patients AA New.*

```

1: /cgn2_6/prodata/1/paa/pct NEW COMB_pep:*
2: /cgn2_6/prodata/1/paa/US06 NEW COMB_pep:*
3: /cgn2_6/prodata/1/paa/US07 NEW COMB_pep:*
4: /cgn2_6/prodata/1/paa/US08 NEW COMB_pep:*
5: /cgn2_6/prodata/1/paa/US09 NEW COMB_pep:*
6: /cgn2_6/prodata/1/paa/US10 NEW COMB_pep:*
7: /cgn2_6/prodata/1/paa/US60 NEW COMB_pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	70	64.8	230	5 US-09-724-676-80619	Sequence 80619, A
2	70	64.8	230	5 US-09-724-676-80620	Sequence 80620, A
3	70	64.8	230	5 US-09-724-676-80621	Sequence 80621, A
4	70	64.8	230	5 US-09-724-676-80622	Sequence 80622, A
5	70	64.8	230	5 US-09-724-676-80624	Sequence 80624, A
6	70	64.8	230	5 US-09-724-676-80625	Sequence 80625, A
7	70	64.8	230	5 US-09-724-676-80626	Sequence 80626, A
8	70	64.8	230	5 US-09-724-676-80627	Sequence 80627, A
9	70	64.8	230	5 US-09-724-676-80628	Sequence 80628, A
10	70	64.8	230	5 US-09-724-676-80629	Sequence 80629, A
11	70	64.8	230	5 US-09-724-676-80630	Sequence 80630, A
12	70	64.8	230	5 US-09-724-676-80631	Sequence 80631, A
13	70	64.8	230	5 US-09-724-676-80632	Sequence 80632, A
14	70	64.8	250	5 US-09-724-676-80588	Sequence 80588, A
15	70	64.8	250	5 US-09-724-676-80589	Sequence 80589, A
16	70	64.8	250	5 US-09-724-676-80590	Sequence 80590, A
17	70	64.8	250	5 US-09-724-676-80591	Sequence 80591, A
18	70	64.8	250	5 US-09-724-676-80592	Sequence 80592, A
19	70	64.8	250	5 US-09-724-676-80593	Sequence 80593, A
20	70	64.8	250	5 US-09-724-676-80594	Sequence 80594, A
21	70	64.8	250	5 US-09-724-676-80595	Sequence 80595, A
22	70	64.8	250	5 US-09-724-676-80597	Sequence 80597, A
23	70	64.8	250	5 US-09-724-676-80598	Sequence 80598, A
24	70	64.8	250	5 US-09-724-676-80599	Sequence 80599, A
25	70	64.8	250	5 US-09-724-676-80600	Sequence 80600, A
26	70	64.8	250	5 US-09-724-676-80601	Sequence 80601, A

RESULT 1

US-09-724-676-80619

; Sequence 80619, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724, 676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO: 80619

; LENGTH: 230

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match Similarity 64.8%; Score 70; DB 5; Length 230; Matches 12; Conservative 63.2%; Pred. No. 0.00045; Indels 0; Gaps 0;

US-09-724-676-80619

RESULT 2

US-09-724-676-80620

; Sequence 80620, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724, 676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO: 80620

; LENGTH: 230

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match Similarity 64.8%; Score 70; DB 5; Length 230; Matches 12; Conservative 63.2%; Mismatches 3; Indels 0; Gaps 0;

US-09-724-676-80620

Db 115 LPRSYDWRKGYYTPVKHQ 133
RESULT 3
 ; Sequence 80621, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 80621
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676-80621

Query Match 64.8%; Score 70; DB 5; Length 230;
 Best Local Similarity 63.2%; Prd. No. 0.00045;
 Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKHQ 19
 Db 115 LPRSYDWRKGYYTPVKHQ 133

Query Match 64.8%; Score 70; DB 5; Length 230;
 Best Local Similarity 63.2%; Prd. No. 0.00045;
 Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKHQ 19
 Db 115 LPRSYDWRKGYYTPVKHQ 133

RESULT 6
 US-09-724-676-80625
 ; Sequence 80625, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 80625
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676-80625

Query Match 64.8%; Score 70; DB 5; Length 230;
 Best Local Similarity 63.2%; Prd. No. 0.00045;
 Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKHQ 19
 Db 115 LPRSYDWRKGYYTPVKHQ 133

RESULT 7
 US-09-724-676-80626
 ; Sequence 80626, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 80626
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676-80626

Query Match 64.8%; Score 70; DB 5; Length 230;
 Best Local Similarity 63.2%; Prd. No. 0.00045;
 Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VPOSIDWRYGAVTSVKHQ 19
 Db 115 LPRSYDWRKGYYTPVKHQ 133

RESULT 8
 US-09-724-676-80627
 ; Sequence 80627, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 80627
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676-80624

Query Match 64.8%; Score 70; DB 5; Length 230;

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 91222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80630
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-80627

Query Match 64.8%; Score 70; DB 5; Length 230;
Best Local Similarity 63.2%; Pred. No. 0.00045; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VPOSIDWRDYGAVTSVKQ 19
Db 115 LPRSTDWREKGYVTPVKQ 133

RESULT 9
US-09-724-676-80628
; Sequence 80628, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80628
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-80628

Query Match 64.8%; Score 70; DB 5; Length 230;
Best Local Similarity 63.2%; Pred. No. 0.00045; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VPOSIDWRDYGAVTSVKQ 19
Db 115 LPRSTDWREKGYVTPVKQ 133

RESULT 10
US-09-724-676-80629
; Sequence 80629, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80629
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-80629

Query Match 64.8%; Score 70; DB 5; Length 230;
Best Local Similarity 63.2%; Pred. No. 0.00045; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VPOSIDWRDYGAVTSVKQ 19
Db 115 LPRSTDWREKGYVTPVKQ 133

RESULT 11
US-09-724-676-80630
; Sequence 80630, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676

RESULT 12
US-09-724-676-80631
; Sequence 80631, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80631
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-80631

Query Match 64.8%; Score 70; DB 5; Length 230;
Best Local Similarity 63.2%; Pred. No. 0.00045; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VPOSIDWRDYGAVTSVKQ 19
Db 115 LPRSTDWREKGYVTPVKQ 133

RESULT 13
US-09-724-676-80632
; Sequence 80632, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80632
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-80632

Query Match 64.8%; Score 70; DB 5; Length 230;
Best Local Similarity 63.2%; Pred. No. 0.00045; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VPOSIDWRDYGAVTSVKQ 19
Db 115 LPRSTDWREKGYVTPVKQ 133

RESULT 14
US-09-724-676-80588
; Sequence 80588, Application US/09724676

```

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 9722
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80588
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-80588

Query Match 64.8%; Score 70; DB 5; Length 250;
Best Local Similarity 63.2%; Pred. No. 0.00049;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VPOSIDWRYGAVTSVKRNQ 19
Db 115 LPNSVDMREKGIVTPVRNQ 133

RESULT 15

US-09-724-676-80589
; Sequence 80589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 9722
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 80589
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-80588

Query Match 64.8%; Score 70; DB 5; Length 250;
Best Local Similarity 63.2%; Pred. No. 0.00049;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VPOSIDWRYGAVTSVKRNQ 19
Db 115 LPNSVDMREKGIVTPVRNQ 133

Search completed: November 21, 2002, 16:16:04

Job time : 8 secs

GenCore version 5.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.
OM protein - protein search, using sw model
Run on: November 21, 2002, 16:10:04 ; Search time 7 Seconds
 (without alignments)
 44.747 Million cell updates/sec

Title: US-09-674-738-1
Perfect score: 108
Sequence: 1. VHQSIDWRDYGAVTISKVKNQ 20
Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 100480 seqs, 15661496 residues

Total number of hits satisfying chosen parameters: 100480

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```

1: /cggn2_6/ptodata/2/pupbaa/US08 NEW PUB.PEP:*
2: /cggn2_6/ptodata/2/pupbaa/PCT NEW PUB.PEP:*
3: /cggn2_6/ptodata/2/pupbaa/US06 NEW PUB.PEP:*
4: /cggn2_6/ptodata/2/pupbaa/US07 PUB.COMB.PEP:*
5: /cggn2_6/ptodata/2/pupbaa/US07 PUB.COMB.PEP:*
6: /cggn2_6/ptodata/2/pupbaa/US07 PUB.COMB.PEP:*
7: /cggn2_6/ptodata/2/pupbaa/US08 PUB.COMB.PEP:*
8: /cggn2_6/ptodata/2/pupbaa/US09 NEW PUB.PEP:*
9: /cggn2_6/ptodata/2/pupbaa/US09 PUB.COMB.PEP:*
10: /cggn2_6/ptodata/2/pupbaa/US10 NEW PUB.PEP:*
11: /cggn2_6/ptodata/2/pupbaa/US10 PUB.COMB.PEP:*
12: /cggn2_6/ptodata/2/pupbaa/US60 NEW PUB.PEP:*
13: /cggn2_6/ptodata/2/pupbaa/US60 PUB.COMB.PEP:*
14: /cggn2_6/ptodata/2/pupbaa/US60 PUB.COMB.PEP:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	21	10	US-09-750-210A-3
2	97	89.8	20	10	US-09-750-210A-5
3	93	86.1	20	10	US-09-750-210A-2
4	88	81.5	18	10	US-09-750-210A-1
5	71	65.7	21	10	US-09-750-210A-1
6	71	65.7	220	10	US-09-462-846-3
7	69	63.9	333	10	US-09-953-956-9
8	69	63.9	333	12	US-10-114-464-9
9	67	62.0	334	10	US-09-529-033-1
10	64	59.3	331	9	US-09-990-064-2
11	64	59.3	331	10	US-09-953-956-8
12	64	59.3	331	12	US-10-114-464-7
13	57	52.8	329	10	US-09-953-956-7
14	57	52.8	329	12	US-10-114-464-7
15	56	51.9	329	12	US-09-953-956-2
16	56	51.9	361	10	US-09-925-300-1693
17	56	51.9	337	10	US-09-925-301-1021
18	55	50.9	335	10	US-09-953-956-10

ALIGNMENTS

RESULT 1
 US-09-750-210A-3
 ; Sequence 3, Application US/09750210A
 ; Patent No. US20020102253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Myntt, Tracy Leanne
 ; APPLICANT: Bogwerda, Christian
 ; APPLICANT: Peak, Keith
 ; TITLE OF INVENTION: Component of Bromelain
 ; CURRENT APPLICATION NUMBER: US/09/750,210A
 ; CURRENT FILING DATE: 2000-12-29
 ; PRIOR APPLICATION NUMBER: US 09/382,688
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: PCT/GB98/00590
 ; PRIOR APPLICATION NUMBER: GB 9706119,6
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: GB 9704252,7
 ; PRIOR FILING DATE: 1997-02-28
 ; PRIOR APPLICATION NUMBER: GB 9703827,7
 ; PRIOR FILING DATE: 1997-02-25
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 3
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: ORGANISM: Ananas comosus
 ; US-09-750-210A-3

Query Match 100.0%; Score 108; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.9e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VHQSIDWRDYGAVTISKVKNQ	20
Db	2	VHQSIDWRDYGAVTISKVKNQ	21

RESULT 2

US-09-750-210A-5 Sequence 5, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Myntott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; CURRENT FILING DATE: 1999-08-25
; CURRENT APPLICATION NUMBER: US/09/750,210A
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 2000-12-29
; FILE REFERENCE: 0023.0740001
; PRIOR FILING DATE: 1998-02-28
; PRIOR APPLICATION NUMBER: GB 9706119, 6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252, 7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850, 9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827, 7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus

US-09-750-210A-5

Query Match 89.8%; Score 97; DB 10; Length 20;
Best Local Similarity 94.7%; Pred. No. 4.e-10;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRYGAUTSVKRNQ 1.9
Db 1 VPOSIDWDRYGAUTSVKRNQ 1.9

RESULT 3
US-09-750-210A-4 Sequence 4, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Myntott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0023.0740001
; CURRENT FILING DATE: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: GB 9704252, 7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850, 9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827, 7
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus

US-09-750-210A-2

Query Match 81.5%; Score 88; DB 10; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.2e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRYGAUTSVKRN 1.8
Db 1 VPOSIDWDRYGAUTSVKRN 1.8

RESULT 4
US-09-750-210A-2 Sequence 2, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Myntott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0023.0740001
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119, 6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252, 7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850, 9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827, 7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus

US-09-750-210A-2

Query Match 81.5%; Score 88; DB 10; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.2e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRYGAUTSVKRN 1.8
Db 1 VPOSIDWDRYGAUTSVKRN 1.8

RESULT 5
US-09-750-210A-1 Sequence 1, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Myntott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peak, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0023.0740001
; CURRENT FILING DATE: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9706119, 6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252, 7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850, 9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827, 7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus

US-09-750-210A-4

Query Match 86.1%; Score 93; DB 10; Length 20;

NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: ORGANISM: Ananas comosus
; US-09-750-210A-1

Query Match Similarity 65.7%; Score 71; DB 10; Length 21;
; Best Local Similarity 68.4%; Pred. No. 9; Mismatches 4; Indels 0; Gaps 0;
; Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RESULT 6
US-09-462-846-3
; Sequence 3, Application US/09462846
; Patent No. US20020311807A1

GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
FILE REFERENCE: GC381-US

CURRENT APPLICATION NUMBER: US/09/462,846
CURRENT FILING DATE: 2000-01-13
PRIORITY NUMBER: PCT/US98/19529
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: EP 97305227.7
PRIOR FILING DATE: 1997-07-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 220
TYPE: PRT

ORGANISM: Bacillus subtilis
; US-09-462-846-3

Query Match Similarity 65.7%; Score 71; DB 10; Length 220;
; Best Local Similarity 63.2%; Pred. No. 0.00012; Mismatches 4; Indels 0; Gaps 0;
; Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKHQ 19
Db 2 IPPSIDWDRQKGAVTEVKNR 20

RESULT 7
US-09-953-956-9
; Sequence 9, Application US/09953956
; Patent No. US20020072107A1

GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHESPIN
FILE REFERENCE: PR07D21

CURRENT APPLICATION NUMBER: US/09/953,956
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: 09/219,441
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
; US-09-953-956-9

Query Match Similarity 63.9%; Score 69; DB 10; Length 333;
; Best Local Similarity 66.7%; Pred. No. 0.00041; Mismatches 3; Indels 0; Gaps 0;
; Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDWDRDGAVTSVKHQ 19
Db 114 IPRSKVDRRKGGYVTPVKHQ 132

RESULT 8
US-10-114-464-9
; Sequence 9, Application US/10114464
; Patent No. US20020142448A1

GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
FILE REFERENCE: PF117D5

CURRENT APPLICATION NUMBER: US/10/114, 464
CURRENT FILING DATE: 2002-04-03
PRIORITY NUMBER: 08/553, 125
PRIOR APPLICATION NUMBER: 08/208, 007
PRIOR FILING DATE: 1994-03-08
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 333
TYPE: PRT

ORGANISM: Homo sapiens
; US-10-114-464-9

Query Match Similarity 63.9%; Score 69; DB 12; Length 333;
; Best Local Similarity 66.7%; Pred. No. 0.00041; Mismatches 3; Indels 0; Gaps 0;
; Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 POSIDWDRDGAVTSVKHQ 19
Db 115 PRSVDRWREKGYVTPVKHQ 132

RESULT 9
US-09-529-063-1
; Sequence 1, Application US/09529063
; Patent No. US20020102542A1

GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAIRICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE POLYPEPTIDE, AND USE OF
; THE BOTH
FILE REFERENCE: 058769
CURRENT APPLICATION NUMBER: US/09/529, 063
CURRENT FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: PCT/JP98/0414
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: JP 9-274674
PRIOR FILING DATE: 1997-10-07
NUMBER OF SEQ ID NOS: 117
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 334
TYPE: PRT
ORGANISM: Homo sapiens
; US-09-529-063-1

Query Match Similarity 62.0%; Score 67; DB 10; Length 334;
; Best Local Similarity 63.2%; Pred. No. 0.00089; Mismatches 4; Indels 0; Gaps 0;
; Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VPOSIDWDRDGAVTSVKHQ 19
Db 114 IPRSKVDRRKGGYVTPVKHQ 132

RESULT 10
US-09-990-064-2
; Sequence 2, Application US/09990064
; Patent No. US20020164765A1

; GENERAL INFORMATION:
; APPLICANT: Parvinder K. Punia
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: CP-30242
; CURRENT APPLICATION NUMBER: US/09/990 , 064
; CURRENT FILING DATE: 001-11-21
; PRIOR APPLICATION NUMBER: GB028462 , 0
; PRIORITY FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mammalia
; US-09-990-064-2

Query Match 59.3%; Score 64; DB 9; Length 331;
Best Local Similarity 57.9%; Pred. No. 0.0027; 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDYGAVTSVKNQ 19
Db 115 LPDSVWREPGCIVTEVKYQ 133

RESULT 13
US-09-953-956-7
; Sequence 7, Application US/09953956
; Patent No. US2002072107A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHESPIN
; FILE REFERENCE: PFI07D2D1
; CURRENT APPLICATION NUMBER: US/09/953 , 956
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 09/219,441
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1

Query Match 52.8%; Score 57; DB 10; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.039; 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 POSIDWRDYGAVTSVKNQ 9
Db 116 PDSIDRKKGVYTPVKNQ 133

RESULT 14
US-10-114-464-7
; Sequence 7, Application US/10114464
; Patent No. US2002042448A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHESPIN
; FILE REFERENCE: PFI07D5
; CURRENT APPLICATION NUMBER: US/10/114 , 464
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/553,125
; PRIOR FILING DATE: 1995-11-07
; PRIOR APPLICATION NUMBER: 08/208,007
; PRIOR FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1

Query Match 52.8%; Score 57; DB 12; Length 329;
Best Local Similarity 66.7%; Pred. No. 0.039; 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 POSIDWRDYGAVTSVKNQ 9
Db 116 PDSIDRKKGVYTPVKNQ 133

RESULT 15

us-9-674-738-1.rapb

; US-09-953-956-2
; Sequence 2, Application US/09953956
; Patent No. US20030072107A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
FILE REFERENCE: FP107D2D1
CURRENT APPLICATION NUMBER: US/09/953, 956
CURRENT FILING DATE: 2001-09-18
PRIORITY APPLICATION NUMBER: 09/219, 441
PRIORITY FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 329
TYPE: PRT
ORGANISM: Homo sapiens
US-09-953-956-2



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Om protein - protein search, using sw model

Run on: November 21, 2002, 16:10:04 ; Search time 7 Seconds

(without alignments) 44.747 Million cell updates/sec

Title: US-09-674-738-2

Perfect score: 105

Sequence: 1 VHQSIQWDRSGAVTSPKNG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100480 seqs, 15661496 residues

Total number of hits satisfying chosen parameters: 100480

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_1 NEW PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCFTS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	105	100.0	20 10 US-09-750-210A-4	Sequence 4, Appli
2	94	89.5	20 10 US-09-750-210A-5	Sequence 5, Appli
3	93	88.6	21 10 US-09-750-210A-3	Sequence 3, Appli
4	79	75.2	18 10 US-09-750-210A-2	Sequence 2, Appli
5	79	75.2	21 10 US-09-750-210A-1	Sequence 1, Appli
6	79	75.2	220 10 US-09-462-846-3	Sequence 3, Appli
7	77	73.3	333 10 US-09-953-956-9	Sequence 9, Appli
8	77	73.3	333 12 US-10-114-464-9	Sequence 9, Appli
9	72	68.6	331 10 US-09-953-956-8	Sequence 8, Appli
10	72	68.6	331 12 US-10-114-464-8	Sequence 8, Appli
11	71	67.6	331 9 US-09-990-064-2	Sequence 2, Appli
12	69	65.7	334 10 US-09-529-063-1	Sequence 1, Appli
13	65	61.9	329 10 US-09-953-956-7	Sequence 3, Appli
14	65	61.9	329 12 US-10-114-464-7	Sequence 9, Appli
15	64	61.6	329 10 US-09-956-2	Sequence 2, Appli
16	64	61.0	329 12 US-10-114-464-2	Sequence 2, Appli
17	64	61.0	361 10 US-09-921-300-1693	Sequence 1693, Appli
18	57.5	54.8	335 12 US-09-956-10	Sequence 10, Appli
19	57.5	54.8	335 12 US-10-114-464-10	RESULT 2

20	57.5	54.8	364 10 US-09-925-302-478	Sequence 478, APP
21	57	54.3	337 10 US-09-825-301-1021	Sequence 1021, APP
22	45.5	43.3	165 10 US-09-864-161-42873	Sequence 42873, APP
23	44	41.9	565 10 US-09-925-301-1302	Sequence 1302, APP
24	43	41.0	119 10 US-09-820-893-50	Sequence 50, APP
25	43	41.0	1247 10 US-09-815-242-1045	Sequence 1045, APP
26	43	41.0	1247 10 US-09-815-242-13841	Sequence 13841, APP
27	42	40.0	185 10 US-09-893-777-78	Sequence 78, APP
28	42	40.0	343 10 US-09-741-669-412	Sequence 412, APP
29	42	40.0	574 10 US-09-782-980-15	Sequence 15, APP
30	42	40.0	574 10 US-09-909-743-6	Sequence 6, APP
31	42	40.0	1165 12 US-10-016-768-10	Sequence 10, APP
32	42	40.0	1221 12 US-10-016-768-11	Sequence 11, APP
33	41	39.0	124 10 US-09-738-973-100	Sequence 100, APP
34	41	39.0	127 10 US-09-728-973-101	Sequence 101, APP
35	41	39.0	190 10 US-09-811-284-227	Sequence 227, APP
36	41	39.0	457 10 US-09-815-242-10370	Sequence 10370, APP
37	41	39.0	516 10 US-09-986-676A-2	Sequence 2, Appli
38	41	39.0	516 10 US-09-911-611-2	Sequence 130, APP
39	40.5	38.6	472 10 US-09-800-729-130	Sequence 95, Appli
40	40.5	38.6	626 10 US-09-800-729-95	Sequence 117, APP
41	39.5	37.6	806 10 US-09-922-217-1117	Sequence 118, APP
42	39.5	37.6	806 10 US-09-922-217-1118	Sequence 119, APP
43	39	37.1	423 10 US-09-829-674-160	Sequence 13686, APP
44	39	37.1	637 10 US-09-815-242-13686	Sequence 13686, APP
45	39	37.1	637 10 US-09-862-207-2	Sequence 2, Appli

ALIGNMENTS

Query Match	100.0%	Score 105; DB 10;	Length 20;	Sequence 478, APP
Best Local Similarity	100.0%	Pred. No. 1.8e-11;		Sequence 1021, APP
Matches 20;	Conervative 0;	Mismatches 0;	Indels 0;	Sequence 1302, APP
SEQ ID NO 4				
LENGTH: 20				
TYPE: PRT				
ORGANISM: ORGANISM:	Ananas comosus			
US-09-750-210A-4				

US-09-750-210A-5 Application US/09750210A
; Sequence 5, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Mynott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623.0740001
; CURRENT APPLICATION NUMBER: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-5

Query Match 89 5%; Score 94; DB 10; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.2e-05; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSATVSKNQ 20
Db 1 VPOSIDWRNYGAVTSVKNQ 20

RESULT 3
US-09-750-210A-3 Application US/09750210A
; Sequence 3, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Mynott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623.0740001
; CURRENT APPLICATION NUMBER: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-3

Query Match 88.6%; Score 93; DB 10; Length 21;

Best Local Similarity 94.7%; Pred. No. 1.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSATVSKNQ 19
Db 2 VPOSIDWRDYGAVTSVKNQ 20

RESULT 4
US-09-750-210A-2
; Sequence 2, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Mynott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623.0740001
; CURRENT APPLICATION NUMBER: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9704252.7
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-2

Query Match 75.2%; Score 79; DB 10; Length 18;
Best Local Similarity 83.3%; Pred. No. 3.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPOSIDWRDGSATVSKN 18
Db 1 VPOSIDWRDYGAVNEVKN 18

RESULT 5
US-09-750-210A-1
; Sequence 1, Application US/09750210A
; Patent No. US20020102253A1
; GENERAL INFORMATION:
; APPLICANT: Mynott, Tracy Lehanne
; APPLICANT: Engwerda, Christian
; APPLICANT: Peek, Keith
; TITLE OF INVENTION: Component of Bromelain
; FILE REFERENCE: 0623.0740001
; CURRENT APPLICATION NUMBER: US/09/750,210A
; CURRENT FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 09/382,688
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: GB 9706119.6
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: PCT/GB98/00590
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703850.9
; PRIOR FILING DATE: 1997-02-25
; PRIOR APPLICATION NUMBER: GB 9703827.7
; PRIOR FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-3

Query Match 88.6%; Score 93; DB 10; Length 21;

```

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 21
; ORGANISM: PRT
; ORGANISM: Ananas comosus
US-09-750-210A-1

RESULT 6
US-09-162-846-3
; Sequence 3, Application US/09462846
; Patent No. US20020031807A1
; GENERAL INFORMATION:
; APPLICANT: Etell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC381-US
; CURRENT APPLICATION NUMBER: US/09/462,846
; CURRENT FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: -PCT/US98/19529
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305227.7
; PRIOR FILING DATE: 1997-07-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-462-846-3

Query Match 75.2%; Score 79; DB 10; Length 21;
Best Local Similarity 70.0%; Pred. No. 3.9e-07;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 2 LPDSIDWRQKGVIEVKNRG 21

RESULT 7
US-09-953-956-9
; Sequence 9, Application US/09953956
; Patent No. US20020072107A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
; FILE REFERENCE: PI07D2D1
; CURRENT APPLICATION NUMBER: US/09/953, 956
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 09/219, 441
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-953-956-8

Query Match 73.3%; Score 77; DB 12; Length 333;
Best Local Similarity 68.4%; Pred. No. 1.9e-05;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 9 IPBIVDWRQKGVITPVKNQG 28

RESULT 8
US-10-114-464-9
; Sequence 9, Application US/10114464
; Patent No. US20020142498A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
; FILE REFERENCE: PI10D5
; CURRENT APPLICATION NUMBER: US/10/114, 464
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/553, 125
; PRIOR FILING DATE: 1995-11-07
; PRIOR APPLICATION NUMBER: 08/208, 007
; PRIOR FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-464-9

Query Match 73.3%; Score 77; DB 10; Length 331;
Best Local Similarity 68.6%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Db 115 LPDSVDRKGCVTEVKYQG 134

RESULT 10
US-10-114-464-8
; Sequence 8, Application US/10114464
; Patent No. US20020142498A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
; FILE REFERENCE: PI07D5
; CURRENT APPLICATION NUMBER: US/10/114, 464

```


US-09-953-956-2
; Sequence 2, Application US/09953956
; Patent No. US20020072107A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, et al.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHPSIN
; FILE REFERENCE: PFI07D2D1
; CURRENT APPLICATION NUMBER: US/09/953, 956
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 09/219, 441
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-953-956-2

Query Match Similarity 61.0%; Score 64; DB 10; Length 329;
Best Local Similarity 63.2%; Pred. No. 0.0027; Mismatches 5; Indels 0; Gaps 0;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 POS1IDRDSGAVTISVENQ 20
| : | || | || || || |
Db 116 PDSVDVKKGYVTPVKNG 134

Search completed: November 21, 2002, 16:16:26
Job time : 8 secs

47A2